09/970944

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

B1

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number
WO 00/73328 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/EP00/05108

(22) International Filing Date: 2 June 2000 (02.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9912755.7 1 June 1999 (01.06.1999) GB

(71) Applicant (for all designated States except US): DEVGEN NV [BE/BE]; Technologiepark 9, B-9052 Zwijnaarde (BE).

(72) Inventors; and
(75) Inventors/Applicants (for US only): VAN CRIEKINGE, Wim [BE/BE]; Devgen NV, Technologiepark 9, B-9052 Zwijnaarde (BE). ROELENS, Ingele [BE/BE]; Devgen NV, Technologiepark 9, B-9052 Zwijnaarde (BE). BOGAERT, Thierry [BE/BE]; Devgen N.V., Technologiepark 9, B-9052 Zwijnaarde (BE). VERWAERDE, Phillipe [FR/BE]; Devgen NV, Technologiepark 9, B-9052 Zwijnaarde (BE).

(74) Agent: BAVERSTOCK, Michael, George, Douglas; Boult Wade Tennant, Verulam Gardens, 70 Gray's Innroad, London WC1X 8BT (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *Without international search report and to be republished upon receipt of that report.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 00/73328 A2

(54) Title: UNC-5 CONSTRUCTS AND SCREENING METHODS

(57) Abstract: The invention provides novel splice variants of the human unc-5c cDNA and a novel human unc-5HS1 cDNA sequence. Also provided are assays based on protein-protein interactions between the UNC-5 protein and a variety of different interacting proteins.

## UNC-5 constructs and screening methods

The present invention is concerned with *unc-5*, a conserved animal gene family that encodes proteins implicated in directional cell behaviour. In particular, the invention is concerned with novel splice variants of the human unc-5C cDNA and a novel human unc-5HS1 cDNA sequence. In addition, assays are provided based on protein-protein interactions between the UNC-5 protein and a variety of different interacting proteins.

*Unc-5* is a conserved animal gene family that encodes proteins implicated in directional cell behaviour. The *unc-5* gene of the nematode worm *Caenorhabditis elegans (C. elegans)* is known to be involved in dorsal migration in contrast to *unc-40* which is involved in ventral migrations (Hedgecock et al., Neuron Vol. 2; 61-85, 1990). Both the *unc-5* and *unc-40* genes are associated with the netrin *unc-6*, and all three genes play a dominant role in directional neuronal outgrowth .

The *C. elegans unc-5* gene encodes a 919 amino acid transmembrane receptor with two immunoglobulin and two thrombospondin type I extracellular domains (Leung-Hagesteijn et al., Cell Vol. 71:289-299, 1992). Ectopic overexpression of unc-5 in the *C. elegans* touch neurons resulted in dorsal steering of these, instead of the normal ventral elongation of these neurons (Hamelin et al., Nature, 364:327-330, 1993).

Several vertebrate homologues of *unc-5* have been cloned including the *Rattus norvegicus unc5H1* and *unc5H2* (Leonardo et al., Nature Vol. 386:833-838, 1997), a Mus musculus homologue designated *rcm* (Ackerman et al, Nature Vol. 386:838-842, 1997) and a human homologue *unc5C* (Ackerman et al., Genomics Vol. 52:205-208, 1998).

The intracellular part of the UNC-5 proteins contains a ZO-1 domain. Such domains are known to be involved in tight junction biology. Furthermore UNC-5 proteins contain a death domain. So far this is the only protein found in *C. elegans* that harbors such a death domain. Death domains are involved in the apoptotic process. In this process, caspases play an important role. The human UNC-40 homologue DCC, a protein also known involved in axonal outgrowth, is a caspase-3 substrate (Mehen et al., Nature 395:801-804, 1998).

The present inventors have identified three previously unknown variant unc-5C cDNAs. These variant cDNAs correspond to alternatively spliced unc-5C transcripts.

Accordingly, in a first aspect provides a protein which comprises the amino acid sequence set forth in SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 6 or an amino acid sequence which differs from that shown in SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 6 only in conservative amino acid changes.

Also provided by the invention are nucleic acid sequences which encode the proteins of the invention.

Also provided by the invention are a nucleic acid comprising the sequences of nucleotides set forth in SEQ ID NO: 1, a nucleic acid comprising the sequences of nucleotides set forth in SEQ ID NO: 3 and a nucleic acid comprising the sequences of nucleotides set forth in SEQ ID NO: 5.

The splice variants of human *unc-5C* were cloned by PCR technology. Two primers were developed to amplify the intracellular part of the unc-5C. Human Brain cDNA was used for this purpose. Three new splice variants of human unc-5C were characterized. A schematic representation of these splice variants is given in Figure 5.

The first splice variant (designated unc-5Cb) has

a deletion of an intron in the UP region. The nucleotide sequence of a partial unc-5Cb cDNA is set forth in SEQ ID NO: 1 and the corresponding amino acid sequence is set forth in SEQ ID NO: 2. The splice of this intron results in a UNC-5Cb protein which is considerably shorter than the previously known UNC-5C, as the coding frame is not maintained. This protein is truncated for the DD domain and for the major part of the UP domain.

The second splice variant (designated unc-5Cc) is deleted by an intron in the ZO-1 region, also resulting in a shorter protein than the previously known UNC-5C, as the coding frame is not maintained. The nucleotide sequence of a partial UNC-5Cc cDNA is set forth in SEQ ID NO: 3 and the corresponding amino acid sequence is shown in SEQ ID NO: 4. The resulting protein (UNC-5Cc) is truncated for the DD domain, the UP domain and a part of the ZO-1 domain.

The third splice variant (unc-5C8) is deleted by a small intron in the ZO-1 domain, but the coding frame is maintained. This results in a slightly smaller protein (UNC-5C8), wherein only the amino acid sequence coded by the spliced intron is truncated. The nucleotide sequence of a partial UNC-5C8 cDNA is set forth in SEQ ID NO: 5 and the corresponding amino acid sequence is shown in SEQ ID NO: 6.

The presence of various splice variants of unc-5C in the human brain indicated that the activity of UNC-5C is tightly regulated.

The inventors have also identified a human unc-5 cDNA which shares homology with the *Rattus norvegicus unc-5H1* cDNA.

Accordingly, in a further aspect the invention provides a nucleic acid molecule comprising the sequence of nucleotides set forth in SEQ ID NO: 7.

Whilst performing yeast two hybrid experiments to identify proteins which interact with the human UNC-5C

protein the inventors identified a number of heretofore unknown human cDNAs which encode proteins which interact with human UNC-5C.

Accordingly, the invention further provides a nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 56 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 57, a nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 54 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 55, a nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 61 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 62 and a nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 63 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 64.

The nucleic acid molecules according to the invention may, advantageously, be included in a suitable expression vector to express the proteins encoded therefrom in a suitable host. Incorporation of cloned DNA into a suitable expression vector for subsequent transformation of said cell and subsequent selection of the transformed cells is well known to those skilled in the art as provided in Sambrook *et al.* (1989), molecular cloning, a laboratory manual, Cold Spring Harbour Laboratory Press.

An expression vector according to the invention includes a vector having a nucleic acid according to the invention operably linked to regulatory sequences, such as promoter regions, that are capable of effecting expression of said DNA fragments. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner.

Such vectors may be transformed into a suitable host cell to provide for expression of a protein according to the invention. Thus, in a further aspect, the invention provides a process for preparing proteins according to the invention which comprises cultivating a host cell, transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the protein, and recovering the expressed protein.

The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, and optionally a promoter for the expression of said nucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable markers, such as, for example, an antibiotic resistance.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and to direct an appropriate level of transcription initiation and also translation initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for translation initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or be assembled from the sequences described by methods well known in the art.

Nucleic acid molecules according to the invention may be inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids, including antisense peptide

nucleic acid (PNA), may be produced by synthetic means.

In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in cases which result in a synonymous codon (a different codon specifying the same amino acid residue) due to the degenerate code in conservative amino acid substitutions. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The nucleic acid sequences according to the invention may be produced using recombinant or synthetic techniques, such as for example using PCR which generally involves making a pair of primers, which may be from approximately 10 to 50 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with cDNA, or genomic DNA from a human cell, performing a polymerase chain reaction under conditions which brings about amplification of the desired region, isolating the amplified region or fragment and recovering the amplified DNA. Generally, such techniques are well known in the art, such as described in Sambrook *et al.* (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids according to the invention may carry a revealing label. Suitable labels include radioisotopes such as $^{32}P$ or $^{35}S$, enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

The protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a protein encoded by said molecule and

having conservative amino acid changes. Proteins or polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said proteins or polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% and preferably 95% amino acid homology with the proteins or polypeptides encoded by the nucleic acid molecules according to the invention. The protein according to the invention may be recombinant, synthetic or naturally occurring, but is preferably recombinant.

A further aspect of the invention provides a host cell or organism, transformed or transfected with an expression vector according to the invention. The host cell or organism may advantageously be used in a method of producing protein, which comprises recovering any expressed protein from the host or organism transformed or transfected with the expression vector.

According to a further aspect of the invention there is also provided a transgenic cell, tissue or organism comprising a transgene capable of expressing a protein according to the invention. The term "transgene capable of expressing" as used herein encompasses any suitable nucleic acid sequence which leads to expression of proteins having the same function and/or activity. The transgene, may include, for example, genomic nucleic acid isolated from human cells or synthetic nucleic acid, including DNA integrated into the genome or in an extrachromosomal state. Preferably, the transgene comprises the nucleic acid sequence encoding the proteins according to the invention as described herein, or a functional fragment of said nucleic acid. A functional fragment of said nucleic acid should be taken to mean a

fragment of the gene comprising said nucleic acid coding for the proteins according to the invention or a functional equivalent, derivative or a non-functional derivative such as a dominant negative mutant, or bioprecusor of said proteins.

The protein expressed by said transgenic cell, tissue or organism or a functional equivalent or bioprecusor of said protein also forms part of the present invention. Recombinant proteins may be recovered and purified from host cell cultures by methods known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose, chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxyapatite chromatography and lectin chromatography.

The protein of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the expressed protein may lack the initiating methionine residue as a result of post-translational cleavage. Proteins which have been modified in this way are also included within the scope of the invention.

In a still further aspect the invention provides an antibody capable of specifically binding to a protein according to the invention. Preferably the antibody is capable of specifically binding to a protein comprising the sequence of amino acids set forth in SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 6. An antibody according to the invention may be raised according to standard techniques well known to those skilled in the art by using the protein of the

invention or a fragment or single epitope thereof as the challenging antigen.

A further aspect of the invention comprises a nucleic acid capable of hybridising to the nucleic acids according to the invention, and preferably capable of hybridising to the sequence of nucleotides set forth in SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63 or SEQ ID NO: 64, under high stringency conditions. Conditions of stringency are well known to those skilled in the art.

Stringency of hybridisation as used herein refers to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature (Tm) of the hybrids. Tm can be approximated by the formula:

$$81.5°C+16.6(\log_{10}[Na^+]+0.41\ (\%G\&C)-600/1$$

wherein 1 is the length of the hybrids in nucleotides. Tm decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences according to the invention.

The present invention also advantageously provides oligonucleotides consisting essentially of at least 10 consecutive nucleotides of a nucleic acid according to the invention and preferably from 10 to 50 consecutive nucleotides of a nucleic acid according to the invention, in particular a nucleic acid comprising the sequence of nucleotides shown in SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63 or SEQ ID NO: 64. These oligonucleotides may, advantageously be

used as probes or primers to initiate replication, or the like. Oligonucleotides having a defined sequence may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex formation between the probe and any nucleic acid in the sample.

To address the functional role of UNC-5 within the cell the inventors used the yeast two hybrid method (Fields and Song, Nature 340:245, 1989), a method well known to molecular biologists, both to investigate the ability of UNC-5 to form dimers and to search for proteins that interact with the UNC-5 protein. Using the two hybrid approach the inventors were able to demonstrate that UNC-5 is capable of forming homodimers and identified a number of proteins which interact with the intracellular domains of the *C. elegans* unc-5 or human UNC-5 proteins. These newly identified protein-protein interactions involving UNC-5 may represent important events in cellular signalling, hence compounds which disrupt these interactions may potentially have useful pharmacological properties.

Accordingly, in a further aspect the invention provides a method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

providing a host cell containing a DNA construct comprising a reporter gene operatively linked to a promoter regulated by a transcription factor having a DNA binding domain and an

activating domain;

expressing in said host cell a first hybrid DNA sequence encoding a first fusion protein comprising an UNC-5 protein or a fragment thereof fused in-frame to either the DNA binding domain or the activating domain of the said transcription factor;

expressing in said host cell a second hybrid DNA sequence encoding a second fusion protein comprising an interacting protein or a fragment thereof fused in-frame to either the DNA binding domain or the activating domain of the said transcription factor, such that when the first fusion protein comprises the activation domain of the said transcription factor the second fusion protein comprises the DNA binding domain of the said transcription factor and when the first fusion protein comprises the DNA binding domain of the transcription factor the second fusion protein comprises the activation domain;

contacting the host cell with a sample of the compound under test; and

detecting any binding of the UNC-5 protein or fragment thereof to the interacting protein or fragment thereof by detecting the production of any reporter gene product in the said host cell.

The method of the invention is based upon the standard two hybrid assay well known in the art. Preferably the host cell is a yeast cell. Protocols for performing a yeast two hybrid assay are well known in the art and are given in the Examples included herein.

As would be readily apparent to persons skilled in the art, the assay can be performed in either orientation. That is to say, the assay can be performed using an UNC-5 protein or a fragment thereof

fused to the DNA binding domain of the transcription factor and the interacting protein or fragment thereof fused to the activation domain of the transcription factor or alternatively the assay can be performed using an UNC-5 protein or a fragment thereof fused to the activation domain of the transcription factor and the interacting protein or fragment thereof fused to the DNA binding domain of the transcription factor.

The above-described method based on the classical yeast two hybrid system can be used to screen for compounds that inhibit or enhance the interaction between two proteins. In addition, other systems have been developed to screen for dissociation events, these methods are designated reverse hybrid methods. These systems make use of yeast strains in which the expression of interacting hybrid proteins increases the expression of a counter-selectable marker that is toxic under particular conditions. Under these conditions, dissociation of an interaction provides a selective advantage, thereby facilitating detection: A few growing yeast colonies in which hybrids fail to interact can be identified among millions of non-growing colonies expressing interacting proteins. Several reverse hybrid systems are known in the art.

The first reverse two-hybrid system utilizes a yeast strain, which is resistant to cycloheximide due to the presence of a mutant CYH2 gene. This strain also contains the wild-type CYH2 allele under the transcriptional control of the GAL1 promoter. Expression of the wild-type GAL4 protein is sufficient to restore growth sensitivity to cycloheximide. Growth sensitivity towards cycloheximide is also restored by the co-expression of the avian c-Rel protein and its IκB-α counterpart, p40, as GAL4 fusion proteins. Restoration of growth sensitivity towards cycloheximide requires the association of c-REL and p40 at the GAL1 promoter and correlates with the

ability of the c-REL/p40 interaction to activate expression from the GAL1 promoter (Leanna and Hannink, 1996, NAR 24:3341-3347)

Another reverse hybrid system makes use of the most widely used counter-selectable marker in yeast genetics, URA3, which encodes orotidine-5'-phosphate decarboxylase, an enzyme required for the biosynthesis of uracil. Yeast cells that contain wild-type URA3, either on a plasmid or integrated in the genome, grow on media lacking uracil (URA3+ phenotype). However, the URA3-encoded decarboxylase can also catalyze the conversion of a non-toxic analogue, 5-fluorooritic acid (FOA) into a toxic product, 5-fluoroacil (Boeke et al., 1984, Mol. Gen. Genet. 197:345-346). Hence mutations that prevent an interaction can be selected from large libraries of randomly mutated alleles. Similarly, molecules that dissociate or prevent an interaction could be selected from large libraries of peptides or compounds (Vidal et al., 1996, PNAS 93:10315-10320; Vidal et al., 1996, PNAS 93:10321-10326).

A third reversed yeast two hybrid is based on the use of GAL80 gene as relay gene. GAL80 encodes a protein that binds to and masks the activation domain of a transcriptional activator, such as GAL4. The reporter genes, which will provide the transcriptional read-out (i.e. HIS3 or LACZ), are dependent upon the functional GAL4 for expression. Only when the level of GAL80 masking protein is reduced by interfering with the two-hybrid interaction will Gal4 function as a transcriptional activator, providing a positive transcriptional read-out for molecules that inhibit the two-hybrid protein-protein interaction. An important feature of this reverse two-hybrid system is that the basal level and the half-time of the relay protein, GAL80, can be fine-tuned to provide maximum sensitivity (Powers and Erickson, 1996, WO95/26400).

The invention further provides a method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

providing a transgenic cell or organism expressing a first fusion protein comprising an UNC-5 protein or a fragment thereof fused in-frame to a first genetically encoded fluorophore and a second fusion protein comprising an interacting protein or a fragment thereof fused in-frame to a second genetically encoded fluorophore, the first and second fluorophores being characterised in that the emission spectrum of one of the fluorophores overlaps with the absorption spectrum of the other fluorophore;

measuring the amount of fluorescence emitted from the fluorophore having an emission spectrum which overlaps with the absorption spectrum of the other fluorophore;

exposing the transgenic cell or organism to a compound under test; and

detecting any change in the amount of fluorescence emitted fluorescence emitted from the fluorophore having an emission spectrum which overlaps with the absorption spectrum of the other fluorophore.

This method uses fluorescence energy transfer or FRET, a technique well known in the art for the detection and quantitative measurement of a whole range of specific binding interactions in biological systems, to screen for compounds which modulate the binding of UNC-5 or a fragment thereof to an interacting protein. The general principles of FRET are as follows: one component of a binding pair is labelled with a first fluorophore (hereinafter

referred to as the donor fluorophore) and a second component of the binding pair is labelled with a second fluorophore (hereinafter referred to as the acceptor fluorophore).

It is an essential feature of the FRET technique that the fluorescence emission spectrum of the donor fluorophore overlaps with the absorption spectrum of the acceptor fluorophore, such that when the two components of the binding pair bind to each other, bringing the donor and acceptor fluorophores into close proximity, a proportion of the fluorescent signal emitted by the donor fluorophore (following irradiation with incident radiation of a wavelength absorbed by the donor fluorophore) will be absorbed by the proximal acceptor fluorophore (a process known in the art as fluorescence energy transfer) with the result that a proportion of the fluorescent signal emitted by the donor fluorophore is quenched and, in some instances, that the acceptor fluorophore emits fluorescence. Fluorescence energy transfer will only occur when the donor and acceptor fluorophores are brought into close proximity by the specific binding reaction. Thus, in the presence of a compound which disrupts the specific binding, the amount of quenching is reduced resulting in an increase in the intensity of the fluorescent signal emitted by the donor fluorophore or a fall in the intensity of the signal emitted by the acceptor fluorophore).

The method of the invention is an *in vivo* FRET assay because it is performed in a transgenic host cell or organism. The transgenic cell can be any mammalian cell line, the transgenic organism is preferably *C. elegans*.

The method of the invention uses genetically encoded donor and acceptor fluorophores which can be expressed as fusion proteins fused in frame to the UNC-5 protein and the interacting protein. This can

be readily accomplished by transforming or transfecting the cell or organism with appropriate expression vectors arranged to express the fusion proteins.

In a preferred embodiment the genetically encoded donor and acceptor proteins are variant green fluorescent proteins which exhibit different fluorescent properties and which have suitably overlapping emission/absorption spectra, such as EGFP (enhanced green fluorescent protein) and EBFP (enhanced blue fluorescent protein). As would be readily apparent to persons skilled in the art, the FRET assay can be performed in either orientation. That is to say, the assay can be carried out using UNC-5 fused to the donor fluorophore and the interacting protein fused to the acceptor fluorophore or using UNC-5 fused to the acceptor fluorophore and the interacting protein fused to the donor fluorophore.

The invention further provides a method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

providing a first reaction component comprising a first protein linked to a solid support containing a scintillant and a second reaction component comprising a second protein which has been radioactively labelled, wherein the first and second proteins are an UNC-5 protein or a fragment thereof and an interacting protein or a fragment thereof;

bringing the first and second reaction components into contact in an aqueous solution in the presence of a compound under test; and

detecting binding of the first protein to

the second protein by detecting light emission from the scintillant.

The above method is based on the scintillation proximity assay (SPA™) developed by Amersham and commonly used in automated high throughput screening. In order to perform this assay a first interacting protein (e.g. an UNC-5 protein) must be linked onto a bead containing a scintillant. Linking of the protein to the beads can be carried out in many different ways, including, for example, via biotin-streptavidin affinity binding. Streptavidin-SPA beads are commercially available from Amersham and the interacting protein can easily be biotinylated *in vitro* or expressed as a biotinylated fusion protein using techniques known in the art. The second interacting protein (e.g. a protein known to interact with UNC-5) is labelled with radioactivity. This can be achieved, for example, by synthesising the second interacting protein by *in vitro* translation and incorporating a tritiated precursor amino acid. The SPA™ assay protocol is then as follows:

SPA beads linked to the first interacting protein are incubated for 30 minutes to one hour with a sample containing the radioactively labelled second interacting protein. Upon binding of the two interacting proteins, the radioactivity emitted by the labelled protein is brought into close proximity with the bead containing scintillant and therefore induces light emission from the scintillant. The free labelled protein in sample (non-bound) will not be held in sufficiently close proximity to the beads to induce light emission. Compounds which disrupt the binding of the first and second interacting proteins will cause a decrease in the amount of light emitted during the experiment.

As would be readily apparent to persons skilled

in the art the assay can be carried out using UNC-5 linked to the solid support containing scintillant and a radioactively labelled interacting protein or using an interacting protein linked to the solid support containing scintillant and a radioactively labelled UNC-5.

The invention further provides a method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

coating the wells of a microtiter plate with UNC-5 protein or a fragment thereof;

contacting the UNC-5 protein or fragment thereof with an aqueous solution comprising an interacting protein or a fragment thereof, said interacting protein being labelled with a tag which is directly or indirectly detectable, and a compound under test;

washing to remove the compound under test and any unbound tagged interacting protein; and

detecting complexes of UNC-5 or a fragment thereof bound to the interacting protein or a fragment thereof by directly or indirectly detecting the presence of the tag.

This method of the invention uses an ELISA type approach to screen for compounds which disrupt binding between Unc-5 and a protein known to interact with UNC-5. In these experiments, the wells of a microtiter plate are coated with the UNC-5 protein or fragments thereof. A sample containing both the compound under test and a protein known to interact with UNC-5 (or a fragment of the protein which is still capable of binding to UNC-5) is then added to the wells and the plates are incubated to allow time for specific

binding of UNC-5 to the interacting protein. The interacting protein (or fragment thereof) is labelled with a tag which is directly or indirectly detectable, typically a fluorescent molecule such as GFP, or a tag which is detectable by specific antibody binding, such as a His-tag or GST-tag. Many other tag molecules which are equally suitable for this purpose are known in the art and are available commercially. The wells are then washed to remove the compound and any interacting proteins which remain unbound. Any interacting protein which has become bound to UNC-5 is not removed by the washing step and can be detected via the directly or indirectly detectable tag. If the interacting protein is labelled with a GFP tag, then bound proteins are detected by measuring GFP fluorescence; if the interacting protein is labelled with a His-tag or a GST tag, bound proteins are detected with immunological techniques, using an antibody of the appropriate specificity.

Compounds which disrupt the binding of UNC-5 to the interacting protein will result in more of the protein remaining unbound, hence less protein will be detected after the washing step.

The invention further provides a method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

exposing a cell or organism expressing UNC-5 and overexpressing nucleic acid encoding an interacting protein to the compound under test; and

screening for reversion of the overexpression phenotype of the cell or organism to wild-type.

Over-expression of genes encoding for proteins which interact with UNC-5 in a cell line or in *C. elegans* results in an over-expression phenotype.

Assays to select for compounds that inhibit the interaction of UNC-5 and its interacting proteins can therefore be performed in cell lines or *C. elegans* by exposing cells or worms exhibiting an over-expression phenotype to the compound under test and screening for a 'reduction' of the over-expression phenotype (i.e. screening for a reversion to wild-type).

Over-expression of proteins which interact with unc-5 in *C. elegans* typically results in neuronal outgrowth phenotypes, distal tip cell outgrowth phenotypes, and other aberrant outgrowth of various tissues and cells. These phenotypes can be easily monitored by expressing reporter genes, such as fluorescent proteins in these cells. Reduction of the phenotype induced by the over-expression can then be monitored by visual inspection.

Simple assays have been developed to screen for compounds which cause reversion of the over-expression phenotype in cell lines. As Unc-5 receives signals from the netrins, over-expression of proteins which interact with unc-5 typically causes phenotypic changes in neuronal outgrowth and cell movement. Accordingly, the step of screening for reduction of the over-expression phenotype can be performed using a laminin assay, a netrin response assay and assays using agarose concentration gradients, a boyden chamber or stratified layers (see Gundersen, R. W., Dev. Biol., 1987, 121(2): 423-431; Klostermann, S. and Bonhoeffer, F., 1996, 4: 237-252). In general, these methods are based upon providing attractants or repellants for axonal guidance in a controlled manner. The way the cells react to these attractants and repellants forms the basis of the assay. In the Boyden chamber (upper and lower chambers separated by

a filter barrier) one typically cultivates cells in the upper chamber and measures how the cells grow through the filter. The agarose approach allows the establishment of gradients to which the cells react by forming specific patterns.

The above-listed methods are all based upon novel interactions between an UNC-5 protein and proteins shown to physically interact with the UNC-5 protein. In preferred embodiments, the UNC-5 protein is a *C. elegans* UNC-5 protein or a human UNC-5 protein. Preferably the human UNC-5 protein is UNC-5C or any of the UNC-5C splice variants identified hereinbefore or UNC-5HS1.

The methods of the invention can also be carried out using fragments of the UNC-5 protein which retain the ability to bind to the interacting protein. Preferably the fragment comprises the intracellular portion of the protein. Various sub-domains of the intracellular portion of the protein or combinations thereof can also be used.

As used herein the term "interacting protein" encompasses any protein which has been demonstrated to interact with an UNC-5 protein. The interacting protein can be a second UNC-5 protein as the examples included herein demonstrate the ability of UNC-5 to form homodimers. The interacting protein can also be a protein identified as interacting with UNC-5 in a yeast two hybrid experiment. A list of proteins identified as interacting with *C. elegans* UNC-5 or human UNC-5 in a yeast two hybrid experiment is given in the Example 4, below. Any of these proteins, or fragments thereof which retain a functional UNC-5 binding site, can be used in the methods of the invention in combination with the appropriate UNC-5 protein or a fragment thereof.

As would be readily apparent to persons skilled

in the art, the UNC-5 signalling pathway is highly conserved across species. Hence it is to be expected that for every interacting protein identified in the yeast two hybrid experiments described in the Examples given herein a homologous interacting protein will be found in other species. For example, for every interacting protein found in *C. elegans* to interact with the *C. elegans* unc-5 protein it is expected that a homologous interacting protein will be found in humans and will interact with a human UNC-5 protein, and vice versa for interacting proteins first identified in humans. Accordingly, it is within the scope of the invention to perform the methods described above with "homologous combinations" of UNC-5 proteins and interacting proteins and even with cross-species combinations e.g. *C. elegans* unc-5 and a human interacting proteins, human UNC-5 and a human homologue of an interacting protein identified in *C. elegans; C. elegans* unc-5 and a human homologue of an interacting protein identified in *C. elegans; C. elegans* unc-5 and a human interacting protein etc. Lists of homologues of the *C. elegans* and human interacting proteins identified in the yeast two hybrid study are given in the Examples included herein.

In a still further aspect the invention provides a method of identifying compounds which reduce or inhibit the lethal phenotype associated with the expression of the UNC-5 death domain in yeast, which method comprises:

exposing a yeast cell containing an expression vector comprising nucleic acid encoding an UNC-5 protein or a fragment thereof comprising the death domain to a compound under test;

allowing the yeast cells to grow in the presence of the compound; and

screening for a reduction or inhibition of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast.

The UNC-5 protein used in the method of the invention is preferably a *C. elegans* UNC-5 protein or a human UNC-5 protein. Preferably the human UNC-5 protein is UNC-5C or any of the UNC-5C splice variants identified hereinbefore or UNC-5HS1.

In a still further aspect the invention provides a method of identifying suppressers of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast, which method comprises:

transfecting yeast cells containing an expression vector comprising nucleic acid encoding an UNC-5 protein or a fragment thereof comprising the death domain with a cDNA library cloned in a yeast expression vector;

allowing the transfected yeast cells to grow for one or more cell divisions; and

screening for reduction or inhibition of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast.

Optionally, the method further comprises the steps of:

identifying a transfected yeast cell exhibiting a reduction or inhibition of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast; and

isolating the cDNA clone(s) present in the transfected yeast cell which is/are responsible for conferring reduction or inhibition of the lethal phenotype.

Again, the UNC-5 protein is preferably a *C. elegans* UNC-5 protein or a human UNC-5 protein. Preferably the human UNC-5 protein is UNC-5C or any of the UNC-5C splice variants identified hereinbefore or UNC-5HS1. The cDNA library is preferably a *C. elegans* cDNA library or a human cDNA library.

The invention will be further understood with reference to the following experimental examples, together with the accompanying Figures in which:

Figure 1 shows a sequence alignment of the known human unc-5C cDNA sequence and the three novel alternative splice variants of unc-5C. The region of alignment corresponds to the portion of the cDNA which encodes the intracellular domains of unc-5C.

Figure 2 shows a multiple alignment of unc-5H1 genes. ym97d12 is an EST clone containing a fragment of the unc-5HS1 cDNA, 3D is a fragment of the unc-5HS1 cDNA cloned by PCR in Example 2.

Figure 3 summarises the cloning of human unc-5C variants.

Figure 4 summarises the cloning of human unc-5HS1.

Figure 5 is a schematic representation of the human unc-5C splice variants.

Figure 6 shows an alignment between a fragment of the protein encoded by the cDNA fragment cloned in pYMP6 and the rat neurexim II-alpha-b cDNA.

Figure 7 shows an alignment between a fragment of the

protein encoded by the cDNA fragment cloned in pYMP17 and the mouse mena protein.

Figure 8 is a representation of the vector pGC1037.

Figure 9 is a representation of the vector pGC1003.

Example 1

Cloning of the human unc-5C splice variants.

Splice variants of human unc-5C were cloned, primary with RACE technology.

A 5′ RACE was performed using the 5′ RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GibcoBRL, Merelbeke, Belgium), according the instructions supplied by the manufacturer or with minor modifications thereof. The primers were based on the unc-5 EST ym97d12.

The first strand cDNA synthesis was performed with primer:

GSP1=oGC75: CGTAGCAGGCACTGGCCTCC

PCR of dC-tailed cDNA: was performed with the gene-specific primer:

GSP2=oGC76: GCACTGGCCTCCAGCTGGCAGTAG

and the RACE anchor primer supplied with the 5′ RACE system.

The PCR Program was:

| | |
|---|---|
| Step 1 | 94°C, 2 min |
| Step 2 | 94°C, 30 sec |
| Step 3 | 60°C, 30 sec |
| Step 4 | 72°C, 2 min |
| Repeat steps 2 to 4 for 35 cycles | |
| Step 5 | 72°C, 7 min |
| Step 6 | 4°C |

A nested PCR was performed with gene-specific primer:

GSP3=oGC77: AGTAGAGGTGGGAGGGCGCCTCCTCGCCCAG

and 5' RACE anchor primer

The PCR program was:

Step 1 92°C, 2 min
Step 2 92°C, 1 min
Step 3 68°C, 2 min
Repeat steps 2 and 3 for 35 cycles
Step 4 72°C, 7 min
Step 5 4°C

The resulting RACE products were visualised by electrophoresis on agarose gels, the bands excised and purified with Jetsorb (Genomed, Germany). The RACE products were ligated into plasmids pAS2 and pGEX-5X-3 with T4 DNA ligase (Amersham pharmacia biotech, NJ, USA), or into a TA cloning vector (Invitrogen, Groningen, the Netherlands). Plasmid DNA was purified prior to sequencing using the Qiagen plasmid purification system (Westburg, Leusden, The Netherlands).

Example 2

Cloning of a new human unc-5 gene.

Human Brain Poly A+ RNA was obtained from Clontech, California, USA and first strand cDNA synthesis performed with the Ready To Go T-Primed First-Strand Kit ((Amersham pharmacia biotech, NJ, USA).

Primers were:

for PCR1:

oGC56: CCGGAATTCCATATGTTAATACTGCCCTTCTGCTGCTAA

oGC66: GCGATCTCTGTAGTTGTGGCCTTG

PCR program was:

Step 1 94°C, 1 min
Step 2 53°C, 30 sec
Step 3 72°C, 2 min
Repeat steps 1 to 3 40 times
Step 4 72°C, 7 min

Step 5 4°C

for PCR2

oGC63: GGGAATTCCATATGTTGTTTGTGTATCGGAAGAATCATC

oGC64: ACGCGTCGACTTAATACTGCCCTTCTGCTGCTAAGGAC

oGC65: CCGGAATTCCTTGTTTGTGTATCGGAAGAATCATC

PCR program was:

Step 1 94°C, 5 min

Step 2 92°C, 30 sec

Step 3 55°C, 30 sec

Step 4 72°C, 2 min

Repeat steps 2 to 4 for 25 cycles

Step 5 72°C, 7 min

Step 6 4°C

The resulting PCR products were isolated, cloned and analysed as described in Example 1.

SEQ ID NO: 7 shows the sequence of a PCR product isolated using the above PCR strategy. This PCR product was designated clone 3D. Figure 2 shows an alignment between the *Rattus norvegicus* unc-5H1 cDNA sequence, the sequence of EST ym97d12, the sequence of clone 3D and the sequences of several other PCR products amplified using the above PCR strategy (1G, 1Jrc and 2Brc).

Example 3

Cloning of two of the fragments of UNC-5 for the dimerization experiment.

A PCR amplification was performed with following primers:

UNC5F: GGT GGT CAT ATG GCC ATG GAG TGC TGT AAA CGT GGC AAT TCA AAA AAG

UNC5R: GGC TGC AGG TCG ACG CCC CGG GGC TTA TGG GGA CAC

AAT TTG TGG

Using the cDNA library used in the yeast two hybrid experiment (Example 4) as template.

PCR program was:

| | |
|---|---|
| Step 1 | 94°C, 1 min |
| Step 2 | 53°C, 30 sec |
| Step 3 | 72°C, 2 min |
| Repeat steps 1 to 3 for 25 cycles | |
| Step 4 | 72°C, 7 min |
| Step 5 | 4°C |

The resulting PCR products were isolated and cloned in frame as Ncol/Sall fragments in the vectors pAS2 and pGAD424 supplied by Clontech (Palo Alto, California, USA).

Example 4

Yeast two Hybrid Experiments

To address the functional role of unc-5 the inventors used the yeast two hybrid method (Fields and Song, Nature 340:245, 1989), a method well known to molecular biologists, to search for the proteins that interact with the UNC-5 protein.

The two hybrid method is based on a pair of fusion proteins. The first fusion protein comprises a first of two interacting proteins fused to the transcriptional activation domain of a bipartite yeast transcription factor; the second fusion protein comprises the second of two interacting proteins fused to the DNA binding domain of the bipartite yeast transcription factor. The principle of the method is that if the two domains of the bipartite transcription factor are physically brought together by binding of the first and second interacting proteins then the

resulting complex will be able to activate transcription from a promoter which contains a target binding site for the transcription factor. The two hybrid assay is commonly used to study protein-protein interactions between two known proteins. It can also be used to screen a library of proteins to identify proteins which interact with a given protein. Both of these uses of the two hybrid system are well known to those skilled in the art.

In the present invention, the yeast two hybrid assay was used to identify proteins which interact with *C. elegans* UNC-5 or human UNC-5 as follows: the intracellular part of UNC-5 or parts thereof were cloned in fusion with the DNA-binding domain of the yeast transcription factor GAL4. A cDNA library was cloned into a vector containing the transcriptional activation domain of GAL4. The fusion proteins were then independently expressed together in yeast containing a reporter gene under the transcriptional control of a promoter containing GAL 4 binding sites (typically GAL1 lacZ or GAL1-HIS3).

Methods

(A) Construction of the *C. elegans* library and standard yeast two hybrid experiments.

Construction of *C. elegans* cDNA libraries, and yeast two hybrid experiments with *C. elegans* cDNA were performed as described by Elledge *et al.*, Proc. Natl. Acad. Sci., 1991, 88:1731-1735, or using the Matchmaker™ maker system supplied by Clontech, California, USA according to the protocol supplied by the manufacturer, or by minor modifications of the above-described methods.

(B) A mating yeast two hybrid experiment.

Mating yeast two hybrid experiments were

performed using plasmid pGC1037 (a plasmid map of pGC1037 is shown in Figure 8 and the complete sequence of the plasmid is given in SEQ ID NO: 91) as bait, and a pre-transformed Human Brain MATCHMAKER cDNA library (Clontech, California, USA) according to the protocol supplied by the manufacture, or with minor modifications thereof.

In brief summary, the steps of the method are as follows:-
Inoculate 1 colony containing the bait plasmid into an overnight culture;
Mate the bait culture and the library culture (24 h);
Plate library mating mixtures;
Incubate for at least 8 days;
Streak big colonies onto SD-3 + 5mM AT-plates (+/- Nylon Membrane);
Stain yeast on Nylon membrane;
Prepare yeast DNA from the positives;
Perform restriction digest, if digest is successful perform backtransformation, using positive and negative controls;
Transform positives into MC1061 cells;
Prepare bacterial DNA using Qiagen Plasmid Mini Purification kit, according to the standard Qiagen protocol; and
Perform DNA sequencing.

All positives obtained in the yeast two hybrid screen were assayed for the specificity of the interaction (against empty vector and irrelevant proteins) using the two hybrid system.

<u>(C) Double-stranded RNA inhibition-RNAi cloning isolation and injection.</u>

Double stranded RNA for RNA inhibition experiments was prepared according to the MEGAscript

protocol (Ambion, UK). RNA isolated using this protocol was purified away from contaminants using the RNeasy system from Qiagen (Westburg, the Netherlands), following the instructions for RNA clean-up supplied by the manufacturer. RNA was injected into the nematodes using standard procedures (Methods in Cell biology, Vol 48, Academic Press, 1995).

Results

(A) Auto-activation and dimerization experiments.

In a first series of experiments, the ability of the intracellular domain of *C. elegans* unc-5 or parts thereof to dimerize or to cause auto-activation was tested. Several plasmids were constructed harboring the intracellular domains of unc-5 and parts thereof. Various domains of unc-5, including the membrane proximal part (MMP), the zonula occludens homology domain (ZO-1), the unknown part (UP) and the Death domain (DD) and were cloned in the vectors pAS2 and pGAD424 (Matchmaker, Clontech, CA, USA). The resulting vectors are summarized in Table 1.

Several constructs containing the death domain were found to be either toxic or auto-activating. Furthermore, by performing homo-dimerization experiments, it was found that the intracellular domain of UNC-5C is capable of forming a homo-dimer. Further experiments led to the conclusion that the ZO-1/UP region is probably responsible for the homo-dimerization. Membrane located signal receptors often form homo- or hetero-dimers prior to intracellular signal transduction. Accordingly, it is postulated that dimer formation in UNC-5 could be a critical event in signalling. Based on a knowledge of this dimerization it is possible to develop assays to screen for compounds which disrupt dimer formation and to identify *unc-5* mutants which are unable to dimerize.

The present inventors have found that in humans UNC-5 proteins may be encoded by at least three genes, the homologous genes unc-5C, unc-5HS1, unc-5HS2. As UNC-5 is an important receptor involved in a vast amount of biological processes, it is considered that more functional homologous genes or unc-5 genes may present in the *Homo sapiens* genome. In addition, the expression of the unc-5 gene does not result in the production of a single transcript. The expression of unc-5C locus can result in the production at least 4 isoforms as a result of alternative splicing events. It is possible that the other unc-5 genes will also express splice variants, which may encode different protein isoforms. Any of these unc-5 isoforms may form dimers, analogous to the homo-dimerisation found for *C. elegans* unc-5. Accordingly, assays can also be developed to screen for chemical substances that alter the dimerization of human unc-5 proteins. Compounds identified using such an assay may have pharmacologically useful properties.

(B) Other receptor dimerizations.

It has been suggested that, in addition to UNC-6, UNC-129 also signals to the UNC-5 receptor (Colavita et al., Science 261:706-709). UNC-6 is also known to signal to UNC-40(DCC). UNC-129 belongs to the TGF-β superfamily. TGF-β receptors, including DAF-1 and DAF-4, do not affect axonal guidance. Although new TGF-β receptors may be found that are involved in axonal guidance, it is more likely that the UNC-129 molecule is able to interact with TSP type I domains, which are present in UNC-5. Such interaction between TGF-β molecules and TSP Type I domains has been shown previously (Schultz-Cherry et al., 1994, J. Biol. Chem. 269, 26775). Furthermore UNC-129 is also involved in the UNC-40 pathway.

Recent studies have provided support for the idea

that the UNC-5 receptor induces switching of UNC-40 from attraction to repulsion (Mehlen et al., Nature 395:801-804, 1998). This suggests a linkage of Unc-5 to oncology since Unc-40 is related to vertebrate DCC (deleted in colorectal cancer), which is a candidate tumour-suppressor gene, and encodes a receptor for netrin-1 (UNC-6). The reversal from attraction towards repulsion in growth cone steering with the two receptors UNC-5 and UNC-40 can be explained by hetero-dimerization between UNC-5 and UNC-40. Such switching of function has also been observed in other biological processes. The UNC-40/UNC-5 interaction may function analogously to the Bax/Bcl-2 interaction involved in apoptosis. Bax can be considered as the protein that protects against apoptosis but the relative titre of both Bax and Bcl-2 in a cell may be important in the decision of cell death.

Given that UNC-5 is capable of forming homodimers, it is postulated that UNC-5 is also capable of forming heterodimers with UNC-40. The UNC-5/UNC-40 heterodimers may act as a functional receptor for UNC-6 and UNC-129. Assays to isolate compounds that influence the interaction between UNC-5 and UNC-40, both enhancing and inhibiting this interaction have therefore been developed. These assays are analogous to the assays as described to isolate compounds that influence the formation of the UNC-5 dimers and the assays for compounds that influence the interaction of UNC-5 with its other interacting proteins (see below).

(C) *C. elegans* UNC-5 interacting proteins

The intracellular part of UNC-5 containing the domains MPP, ZO-1 and UP cloned in vector pGC1003 (a plasmid map of pGC1003 is given in Figure 9 and the complete sequence of the plasmid is given in SEQ ID NO: 92) was used as 'bait' in a yeast two hybrid

experiment screening against a *C. elegans* cDNA library. These experiments resulted in the identification of ten genes, including three known genes and seven genes with heretofore unknown function, encoding proteins which specifically interact with the intracellular part of UNC-5. Details of the UNC-5 interacting proteins identified during the two hybrid screen are given below. In most cases, the results of double-stranded RNA inhibition experiments (RNAi) designed to inhibit expression of the interacting protein are also given. Where appropriate, details of human homologues of the interacting protein are also given and any known disease associations are discussed.

**1) Spectrin β-chain / Fodrin β-chain (pC1025)**

A first series of hits resulted in the identification of plasmid pC1025 which contains a fragment of a cDNA encoding the *C. elegans* spectrin β-chain/Fodrin. The spectrin β-chain protein is encoded by the gene K11C4.3, located on chromosome IV.

The full length cDNA and amino acid sequences of spectrin β-chain/Fodrin are shown in SEQ ID NOS: 11 and 12, respectively. The nucleotide sequence of the fragment of the spectrin β-chain cDNA which is cloned as an insert in plasmid pC1025 is given in SEQ ID NO: 13, the corresponding amino acid sequence is given in SEQ ID NO: 14.

RNAi experiments using a double-stranded RNA corresponding to the cDNA fragment cloned in pC1025 revealed that inhibition of the expression of the native spectrin β-chain in *C. elegans* worms causes the following phenotype: no embryonal lethality, normal canals, normal elongation, growth retardation and growth arrest at L1 and L2, nearly no movement but touch reflex is observed. The phenotype is 100% penetrant, and the larvea are short and wrinkled.

These RNAi phenotypes and the corresponding knock-out phenotype can be used as the basis of a compound screen in *C. elegans* to identify chemical substances that modulate the activity of the spectrin β-chain protein.

Human Fodrin (genbank accession number 2493434) contains an extra C-terminal PDZ domain that is not present in spectrin (genbank accession number 134798). The human fodrin seems to be more homologous to the *C. elegans* protein. This is in agreement with the finding that unc-5 is also expressed in the brain of vertebrates.

The interaction between UNC-5 and fodrin could be a critical event in a cell signalling, hence compounds which modulate the interaction between UNC-5 and fodrin, particularly the interaction between human UNC-5 and human fodrin, may potentially have pharmacological activity. Assays can also be developed to screen for genetic mutations that inhibit the interaction needed for proper signal transduction. Compounds which enhance or inhibit the interaction of UNC-5 with fodrin and spectrin β-chain may be useful in the development of pharmaceutical preparations for the treatment of Crohn's disease, Sjogren's syndrome, secretion related diseases, diseases related to neutophil and platelet activation, and long-term potential in neurons, Alzheimer's disease, proliferative diseases such as carcinomas, neoplasia, and more specifically, shwannomas, meningiomas, ependymonas, squamous cell carcinomas, malignant melanomas and lung carcinomas, spherocytosis, pyropoikilocytosis, Duchenne muscular dystrophy and various neurological disorders.

**<u>2) APR-1 (pC1028)</u>**

A second plasmid isolated in the yeast two hybrid

screen, pC1028, contained a fragment of a cDNA encoding APR-1.

The nucleotide sequence of the full length APR-1 cDNA is shown in SEQ ID NO: 15 and the amino acid sequence of the APR-1 protein encoded by this cDNA is shown in SEQ ID NO: 16. The nucleotide sequence of the fragment of the APR-1 cloned in pC1028 is shown in SEQ ID NO: 17, with the corresponding amino acid sequence shown in SEQ ID NO: 18.

RNAi experiments using a double-stranded RNA corresponding to the fragment cloned in pC1028 demonstrated that inhibition of APR-1 expression in *C. elegans* results in the following phenotype: more than 95% embryonic lethality, in 25% of cases this was due to the overproduction of pharyngeal tissue and lack of endoderm, and premature division of the E daughters (Rocheleau *et al.*, Cell 90:707-716, 1997). Escapers (worms that survive) have abnormal gut cells. These RNAi phenotypes and the corresponding knock-out phenotype can be used as the basis of a compound screen in C. elegans to identify chemical entities that modulate the activity of APC (see below), and hence the unc-5 pathway.

Further yeast two hybrid experiments were performed in order to more precisely determine the position of the APR binding regions in UNC5, using the UNC5 domains MPP, ZO-1, UP and combinations thereof. APR-1 seemed to associate with two distinct regions in UNC5. First, APR-1 appears to bind to the MPP domain. Secondly, APR-1 appears to binding to the ZO-1/UP domain. APR-1 seems to bind less to the ZO-1 and UP domains when they are present alone and not in combination. A similar experiment was carried out using the *C. elegans* UNC-5 protein, and domains of human APC and analogous results were obtained. It is concluded that APC is capable of binding to two

distinct regions of UNC-5, the MPP and the ZO-1/UP domains.

The interaction between UNC-5 and APC/APR-1 could be a critical event in cellular signalling and hence compounds which modulate this interaction, particularly compounds which modulate an interaction between human UNC-5 and human APC, may potentially have pharmacological activity. Furthermore genetic mutations, and splice variants can be identified that inhibit the interaction, needed for proper signal transduction. Compounds which enhance or inhibit the interaction of UNC-5 with APR/APC may be useful in the development of pharmaceutical agents for the treatment of neurological diseases and colorectal cancers such as adenomatous polyposis coli.

**<u>3) UNC-14 (pC1034)</u>**

A third plasmid identified during the yeast two hybrid screen using *C. elegans* UNC-5 as bait (pC1034) was found to contain a fragment of the UNC-14 cDNA.

The nucleotide sequence of the full length UNC-14 cDNA is shown in SEQ ID NO: 19, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 20. The nucleotide sequence of the fragment of the UNC-14 cDNA cloned as an insert in pC1034 is shown in SEQ ID NO: 21, with the corresponding amino acid sequence of the polypeptide encoded by this fragment shown in SEQ ID NO: 22.

*C. elegans* worms mutated in unc-14 are observed to be very sluggish, almost paralysed, small, dumpyish, with a tendency to coil and show some egg retention. This phenotype can be used as the basis of a compound screen in *C. elegans* to identify chemical entities that modulate the activity of UNC-14.

Furthermore, *C. elegans* worms mutated in the *unc-14* gene were shown to have abnormal axonal elongation and axonal structures. The unc-14 gene

encodes a protein of 665 amino acids, and is co-expressed with the *unc-51* gene in the cell bodies and axons of almost all neurons including DD/VD and hermaphrodite-specific neurons. The results of yeast two-hybrid experiments suggested that a central region of UNC-14 binds to the carboxy-terminal region of UNC-51, and that the UNC-51 carboxy-terminal region oligomerized (Ogura et al., Genes Dev. 11:1801-1811, 1997).

Mutations in the *unc-51* gene, isolated from mutants of Caenorhabditis elegans exhibiting abnormal axonal extension and growth, encodes a novel serine/threonine kinase (K. Ogura, et al., 1994, Genes Dev. 8: 2389- 2400).

**4) F11A10.1 (pGC1021)**

A fourth plasmid isolated during the yeast two hybrid screen, pGC1021, was found to contain a fragment of cDNA corresponding to the *C. elegans* gene designated F11A10.1.

The nucleotide sequence of the full length F11A10.1 cDNA is shown in SEQ ID NO: 23, with the amino acid sequence of the protein encoded by this cDNA shown in SEQ ID NO: 24. The nucleotide sequence of the fragment of the F11A10.1 cDNA cloned in pGC1021 is shown in SEQ ID NO: 25, the amino acid sequence of the protein fragment encoded by this fragment of the cDNA is shown in SEQ ID NO: 26.

To date, no function is as yet known for F11A10.1. RNAi experiments using a double-stranded RNA corresponding to the insert of pGC1021 showed that inhibition of F11A10.1 expression in *C. elegans* results in worms which are weakly constipated. In *C. elegans*, constipation has been associated with neuronal dysfunction (Thomas, Genetics 124:855-872, 1990). Furthermore and remarkably inhibition of F11A10.1 expression causes migration defects in the

distal tip cell, similar to those observed in unc-5 mutants and unc-14/unc-51 double mutants. These RNAi phenotypes and the corresponding knock-out phenotypes can be used as the basis of a compound screen in C. elegans to identify chemical entities that modulate the activity of F11A10.1.

The interaction between UNC-5 and F11A10.1 could be a critical event in cellular signalling and hence compounds which modulate this interaction may potentially have pharmacological activity. Furthermore, genetic mutations, and splice variants can be identified that inhibit the interaction, needed for proper signal transduction. Compounds which enhance or inhibit the interaction of UNC-5 with F11A10.1 may be of use in the development of pharmaceutical compositions useful in the treatment of neurological disorders, tumours such as Kaposi's Sarcoma, immunological disorders and diseases related to vesicle fusion, proteolysis, peroxisomal and mitochondrial biogenesis, and transcription.

**5) C15E6.1/2 (pGC1026)**

A fifth plasmid identified during the yeast two hybrid experiment, pGC1026, was found to contain a fragment of a cDNA encoding the C15E6.1 protein.

The nucleotide sequence of the full length C15E6.1/2 cDNA is shown in SEQ ID NO: 27, with the amino acid sequence of the protein encoded by this cDNA shown in SEQ ID NO: 28. The nucleotide sequence of the fragment of the C15E6.1/2 cDNA cloned in pGC1026 is shown in SEQ ID NO: 29, the amino acid sequence of the protein fragment encoded by this fragment of the cDNA is shown in SEQ ID NO: 30.

RNAi experiments using a double-stranded RNA corresponding to the insert of pGC1026 did not result in any clear visual phenotype.

The identification of C15E6.1/2 as an UNC-5

interacting protein indicates that UNC-5 might be a band 4 .1 binding protein and may share homology with other band 4.1 binding proteins such as CD44, glycophrin C, and paranodin.

By using the band 4.1 signature to search a database of *C.elegans* genes, F07A11.1 on chromosome II was identified as encoding a band 4.1 protein.

The interaction between UNC-5 and C15E6.1/2 could be a critical event in cellular signalling and hence compounds which modulate this interaction may potentially have pharmacological activity. Furthermore genetic mutations, and splice variants can be identified that inhibit the interaction, needed for proper signal transduction. Compounds which enhance or inhibit the interaction of UNC-5 with C15E6.1/2 may be useful in the development of pharmaceutical preparations for the treatment of diseases related to axonal signalling, synaptic vesicle exocytosis, cell adhesion, cytoskeleton associated proteins, cell morphology , cell growth, allergic inflammatory processes and rheumatoid arthritis.

<u>6) D1081.7 (pGC1027)</u>

A sixth plasmid identified during the two hybrid screen was found to contain a fragment of cDNA corresponding to the *C. elegans* gene designated D1081.7.

The nucleotide sequence of the full length D1081.7 cDNA is shown in SEQ ID NO: 31, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 32. The nucleotide sequence of the fragment of the D1081.7 cDNA cloned as an insert in pGC1027 is shown in SEQ ID NO: 33, with the corresponding amino acid sequence of the polypeptide encoded by this fragment shown in SEQ ID NO: 34.

RNAi experiments performed using double stranded RNA corresponding to the insert in pGC1027 appeared

not to result in any clear visual phenotype.

All genes so far found in *C. elegans* have human homologues. It is therefore expected that D1081.7 will also have vertebrate, including human, homologues. These homologues can be cloned using standard technologies.

The interaction between UNC-5 and D1081.1 could be a critical event in cellular signalling and hence compounds which modulate this interaction may potentially have pharmacological activity and thus be of use in the development of pharmaceutical compositions. Furthermore genetic mutations, and splice variants may be identified that inhibit the interaction needed for proper signal transduction.

**7) B0238.9 (pGC1032)**

A seventh plasmid identified during the two hybrid screen was found to contain a fragment of cDNA corresponding to the *C. elegans* gene designated B0238.9.

The nucleotide sequence of the full length B0238.9 cDNA is shown in SEQ ID NO: 35, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 36. The nucleotide sequence of the fragment of the B0238.9 cDNA cloned as an insert in pGC1032 is shown in SEQ ID NO: 37, with the corresponding amino acid sequence of the polypeptide encoded by this fragment shown in SEQ ID NO: 38.

B0238.9 is located in the chromosomal region where *seu-2* is also located. The *seu-2* was identified in suppressor screens of ectopically expressed *unc-5* and is considered to be involved in the *unc-5* pathway (Colavita and Culotti, Dev. Biol. 194:72-85, 1998). As a gene has now been isolated that interacts with *unc-5*, it is high probable that B0238.9 is the same as *seu-2*. Mutations in *seu-2* appeared not to have any visual phenotype, as was also observed in RNAi

experiments using a double stranded RNA corresponding to a fragment of B0238.9. The finding that SEU-2 is a suppressor and a binding partner to UNC-5 validates the importance of this interaction. Other known suppressors of ectopic *unc-5* growth cone steering are *unc-6, unc-40, unc-34, unc-44, unc-129, seu-1, seu-2,* and *seu-3*. Mutations in some of these genes show axonal guidance defects, unlike *seu-2.*

Homology searches in the EST database with B0238.9 revealed the presence of at least two human ESTs with significant homology. The ESTs so found, nz77b06 and yu53g01, can be used as basis to clone the full length cDNA encoding the human homologue of B0238.9.

The interaction between UNC-5 and B0238.9 could be a critical event in cellular signalling and hence compounds which modulate this interaction may potentially have pharmacological activity and thus may be useful in the development of pharmaceutical compositions. Furthermore genetic mutations, and splice variants may be identified that inhibit the interaction, needed for proper signal transduction.

**8) ZC404.8 (pGC1033)**

An eighth plasmid identified during the two hybrid screen was found to contain a fragment of a cDNA corresponding to the *C. elegans* gene designated ZC404.8.

The nucleotide sequence of the full length ZC404.8 cDNA is shown in SEQ ID NO: 39, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 40. The nucleotide sequence of the fragment of the ZC404.8 cDNA cloned as an insert in pGC1033 is shown in SEQ ID NO: 41, with the corresponding amino acid sequence of the polypeptide encoded by this fragment shown in SEQ ID NO: 42.

RNAi experiments using a double stranded RNA

corresponding to a fragment of this gene resulted in an embryonic lethal phenotype. The worms showed no elongation and only very little muscle activity, the hypodermis is clearly abnormal.

Homology searches in the EST database with ZC404.8.9 revealed the presence of at least three human ESTs with significant homology. The ESTs thus identified, qe69h03, zx61d04, and zd35e10, can be used as basis to clone the full length cDNAs.

The interaction between UNC-5 and ZC404.8 could be a critical event in cellular signalling and hence compounds which modulate this interaction may potentially have pharmacological activity and thus be useful in the development of pharmaceutical preparations. Furthermore genetic mutations, and splice variants may be identified that inhibit the interaction, needed for proper signal transduction.

**9) yk17a3 (pGC1023)**

A ninth plasmid identified during the yeast two hybrid experiment was found to contain a fragment of cDNA corresponding to the *C. elegans* gene designated yk17a3.

The nucleotide sequence of the fragment of the yk17a3 cDNA cloned as an insert in pGC1023 is shown in SEQ ID NO: 43, with the corresponding amino acid sequence of the polypeptide encoded by this fragment shown in SEQ ID NO: 44.

RNAi experiments using a double stranded RNA corresponding to a fragment of yk17a3 resulted in the following phenotypes in *C. elegans*: Very slow growth, and the larvae get typical darker spots as they get older. Inhibition of yk17a3 expression in some non wild-type genetic backgrounds leads to defective moulting, where the worm cannot escape from the old cuticle and therefore shrinks and stays in the L4 stage. The defective moulting phenotype is also

observed when yk17a3 expression is inhibited on a wild-type genetic background, although the phenotype is less prominent. Worms which escape the defective moulting phenotype show defects in vulva development, either lacking a vulva altogether or having a vulva which is non-functional.

Homology searches in the Genbank database with yk17a3 revealed the presence of at least one human homologue of this gene, designated KIAA0187.

The interaction between UNC-5 and yk17a3 (KIAA0187) could be a critical event in cellular signalling and hence compounds which modulate this interaction may potentially have pharmacological activity. Furthermore genetic mutations, and splice variants may be identified that inhibit the interaction, needed for proper signal transduction. Compounds which enhance or inhibit the interaction of UNC-5 with yk17a3 may be of use in the development of pharmaceutical compositions for the treatment of CADASIL, artheriohepatic dysplasia, Alzheimer's disease, neoplasia such as T-cell acute lymphoblastic leukemia and certain cancers, such as pancreatic cancer and colon cancer.

**10) F41H10.3 (pGC1020)**

A tenth plasmid identified using the yeast two hybrid experiment was found to contain a fragment of a cDNA corresponding to the *C. elegans* gene designated F41H10.3.

The nucleotide sequence of the full length F41H10.3 cDNA is shown in SEQ ID NO: 45, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 46. The nucleotide sequence of the fragment of the F41H10.3 cDNA cloned as an insert in pGC1020 is shown in SEQ ID NO: 47, with the corresponding amino acid sequence of the polypeptide encoded by this fragment shown in SEQ ID NO: 48.

F41H10.3 harbors a ATP/GTP binding domain.

Worms resulting from RNAi experiments using a double stranded RNA corresponding to a fragment of F41H10.3 did not exhibit a clear visual phenotype.

All genes so far found in *C. elegans* have human homologues. It is therefore expected that F41H10.3 will also have vertebrate, including human, homologues. These homologues can be cloned using standard technologies well known to persons skilled in the art.

The interaction between UNC-5 and F41H10.3 could be a critical event in signalling and compounds which modulate this interaction may potentially have pharmacological activity and thus be useful in the development of pharmaceutical preparations. Furthermore genetic mutations, and splice variants may be identified that inhibit the interaction, needed for proper signal transduction.

(D) Human UNC-5 interacting proteins.

The intracellular part of the human UNC-5 protein (human UNC-5HS1) containing the domains ZO-1, UP and DD cloned in vector pGC1037 (see above) was used as 'bait' in a yeast two hybrid experiment screening against a pretransformed human brain Matchmaker cDNA library (Clontech, Palo Alto, California USA) using the mating screen approach described above. These experiments resulted in the identification of six genes encoding proteins which interact with UNC-5, including two known genes and four heretofore unknown genes.

All proteins found in this yeast two hybrid screen with the human UNC-5 were different to the proteins found in the screen with the *C. elegans* UNC-5. There are at least two reasons for this variation in the isolated proteins. First, the screens are not saturated, which means that not all possible

interacting proteins have been isolated, neither in the screen with the *C. elegans* UNC-5 nor in the screen with the human UNC-5. Secondly, different intracellular fragments have been used in the screens. In the *C. elegans* UNC-5 screen, the intracellular domains MPP, ZO-1 and UP were used as bait, whereas in the human UNC-5 screen, the intracellular domains ZO-1, UP and DD were used as bait. Proteins with specific interaction patterns will not be isolated if the necessary interacting domain is missing, or if the optimal combination of domains is missing. This has been shown in the *C. elegans* UNC-5 interaction with APR. APR interacts clearly with the MPP domain and the domain combination ZO-1,UP, but interacts less efficiently to with domain combination MPP, ZO-1, although the MPP domain is present. APR binds efficiently to the domain combination MPP, ZO-1, UP.

The human UNC-5 interacting proteins identified during the two hybrid screen are listed below. In each case, any known disease associations are discussed and genes/cDNAs encoding homologous *C. elegans* proteins are listed.

**1) i-beta-1,3-N-acetylaminyltansferase (pYMP5).**

A first plasmid identified during the yeast two hybrid experiment was found to contain a fragment of the cDNA encoding i-beta-1,3-N-acetylaminyltansferase.

The nucleotide sequence of the full length i-beta-1,3-N-acetylaminyltansferase cDNA is shown in SEQ ID NO: 49, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 50. The partial nucleotide sequences of the fragment of the i-beta-1,3-N-acetylaminyltansferase cDNA cloned as an insert in pYMP5 are shown in SEQ ID NOs: 51 and 52, with the corresponding amino acid sequence of the polypeptide encoded by these partial sequences shown in SEQ ID NO: 53.

*C. elegans* has at least seven putative homologues of i-beta-1,3-N-acetylaminyltansferase, designated F22F7.6, C18G1.3, K09C8.4, F21H7.10, C54C8.2, F56H6.6 and T15D6.4. cDNA and/or amino acid sequences for each of these putative homologues are given herein. Amino acid and nucleotide sequences for these homologues are given in SEQ ID NOS: 66 to 82.

The interaction between UNC-5 and beta-1,3-N-acetylglucosaminyltransferase could be a critical event in signalling and hence compounds which modulate this interaction may potentially have pharmacological activity. Furthermore genetic mutations, and splice variants may be identified that inhibit the interaction, needed for proper signal transduction. Compounds which modulate the interaction of UNC-5 with beta-1,3-N-acetylglucosaminyltransferase may be useful in the development of pharmaceutical preparations for the treatment of synaptic cleft dysfunctions, vesicle transport dysfunctions, inflamation, various tumours and more particular in tumour cell adhesion, migration and invasion, such as pancreas cancer, squamous cell cancer, human breast cancer, thyroid neoplasms, colorectal carcinomas.

**2) new gene with slight homology to neurexin II-alpha-b (NHII)(pYMP6)**

A second plasmid identified during the yeast two hybrid experiment was found to contain a fragment of a cDNA corresponding to a new gene with slight homology to neurexin II-alpha-b. The new gene was designated NHII.

Partial nucleotide sequences for the fragment of cDNA cloned as an insert in pYMP6 are shown in SEQ ID NO: 54 (coding strand sequenced from one end of the insert of pYMP6 sequenced with forward primer)and and SEQ ID NO: 55 (non-coding strand sequenced from one end of pYMP6 with reverse primer). The plasmid pYMP6

was deposited in the Belgian Co-ordinated Collections of Microorganisms (BCCM), Universiteit Gent, K. L. Ledeganckstraat 35, B-9000, Gent, Belgium on 21 May 1999 under accession number LMBP 3932. The cDNA insert (approximately 1800bp) can easily be excised from this plasmid by digestion with the restriction enzymes EcoRI and XhoI. Alternatively the cDNA insert sequence can be amplified by PCR using primers corresponding to the sequences for the ends of the insert given in SEQ ID NOS: 54 and 55.

The interaction between UNC-5 and the new gene with homology to neurexin II-alpha-b could be a critical event in signalling and hence compounds which modulate this interaction may potentially have pharmacological activity.

**3) New Gene with Mena homology (MHI) (pYMP17)**

A third plasmid identified during the yeast two hybrid experiment was found to contain a fragment of a cDNA encoding a protein sharing slight homology with the human mena protein. The new gene was designated MHI.

Partial nucleotide sequences of the fragment of cDNA cloned as an insert in pYMP17 are shown in SEQ ID NO: 56,(coding strand sequenced from one end of the insert of pYMP sequenced with forward primer)and SEQ ID NO: 57 (non-coding strand sequenced from one end of pYMP with reverse primer). An alignment between the amino acid sequence encoded by the insert of pYMP17 and the mouse mena protein is shown in Figure 7. The plasmid pYMP17 was deposited in the Belgian Co-ordinated Collections of Microorganisms (BCCM), Universiteit Gent, K. L. Ledeganckstraat 35, B-9000, Gent, Belgium on 21 May 1999 under accession number LMBP 3935. The cDNA insert (approximately 1000bp) can easily be excised from this plasmid by digestion with the restriction enzymes EcoRI and XhoI. Alternatively

the cDNA insert sequence can be amplified by PCR using primers corresponding to the sequences for the ends of the insert given in Figures 55A and 55B.

*C. elegans* has at least one protein with homology to the new Mena homologue (MHI), encoded by the gene designated Y50D4.Contig200. The *C. elegans* gene, unc-34 (which maps with Y50D4) is known to suppress the axonal guidance defects induced by ectopic expression of the Netrin receptor UNC-5 (Colavita, A. et al., Dev.Biol., 194:72-85, 1998.).

The interaction between UNC-5 and mena, members of this mena superfamily, unc-34, and Y50D4.contig200, could be a critical event in signalling and hence compounds which modulate these interactions may potentially have pharmacological activity and thus may be useful in the development of pharmaceutical compositions.

**4) Alpha-2 macroglobulin (pYMP30)**

A fourth plasmid identified during the yeast two hybrid experiment was found to contain a fragment of the human alpha-2 macroglobulin cDNA.

The nucleotide sequence of the full length alpha-2 macroglobulin cDNA is shown in SEQ ID NO: 58, the amino acid sequence of the protein encoded by this cDNA is given in SEQ ID NO: 59. A partial nucleotide sequence for the fragment of the alpha-2 macroglobulin cDNA cloned as an insert in pYMP30 is shown in SEQ ID NO: 60.

*C. elegans* has at least one homologue of alpha-2 macroglobulin, designated ZK337.1, of which two splice variants designated ZK337.1a and ZK337.1b are known to exist.

The interaction between UNC-5 and alpha-2 macroglobulin could be a critical event in signalling and hence compounds which modulate this interaction may potentially have pharmacological activity.

Compounds which enhance or inhibit the interaction of UNC-5 with alpha-2 macroglobulin could be useful in the development of pharmaceutical substances.

**5) New gene 1 (pYMP11)**

A fifth plasmid identified during the yeast two hybrid experiment was found to contain a fragment of cDNA with no homology to any known human cDNA.

Partial nucleotide sequences for the fragment of cDNA cloned as an insert in pYMP11 are shown in SEQ ID NO: 61 (coding strand sequenced from one end of the insert of pYMP sequenced with forward primer)and SEQ ID NO: 62 (non-coding strand sequenced from one end of pYMP with reverse primer). The plasmid pYMP11 was deposited in the Belgian Co-ordinated Collections of Microorganisms (BCCM), Universiteit Gent, K. L. Ledeganckstraat 35, B-9000, Gent, Belgium on 21 May 1999 under accession number LMBP 3933. The cDNA insert (approximately 2300bp) can easily be excised from this plasmid by digestion with the restriction enzymes EcoRI and XhoI. Alternatively the cDNA insert sequence can be amplified by PCR using primers corresponding to the sequences for the ends of the insert given in Figures 59A and 59B.

The interaction between UNC-5 and the protein encoded by the insert of pYMP11 could be a critical event in signalling and hence compounds which modulate this interaction may potentially have pharmacological activity and thus could be useful in the development of pharmaceutical substances.

**6) New gene 2 (pYMP12)**

A sixth plasmid identified during the yeast two hybrid experiment was found to contain a fragment of cDNA with no homology to any known human cDNA.

Partial nucleotide sequences for the fragment of

cDNA cloned as an insert in pYMP12 are shown in SEQ ID NO: 63 (coding strand sequenced from one end of the insert of pYMP sequenced with forward primer)and SEQ ID NO: 64 (non-coding strand sequenced from one end of pYMP with reverse primer). The plasmid pYMP12 was deposited in the Belgian Co-ordinated Collections of Microorganisms (BCCM), Universiteit Gent, K. L. Ledeganckstraat 35, B-9000, Gent, Belgium on 21 May 1999 under accession number LMBP 3934. The cDNA insert (approximately 2000bp) can easily be excised from this plasmid by digestion with the restriction enzymes EcoRI and XhoI. Alternatively the cDNA insert sequence can be amplified by PCR using primers corresponding to the sequences for the ends of the insert given in Figures 60A and 60B.

The interaction between UNC-5 and the protein encoded by the insert of pYMP12 could be a critical event in signalling and hence compounds which modulate this interaction may potentially have pharmacological activity and thus could be useful in the development of pharmaceutical substances.

Example 5

Yeast two hybrid compound screens

Interactions of proteins leads to expression of a reporter protein β-galactosidase in a yeast two hybrid assay. An assay has been developed that is usable in 96 or 384 well plates or microtiter plates with another number of wells. This assay is suitable for high throughput compound screening. Optimal performance of the assay is dependent upon at least two important parameters:
lysis of yeast cells and the choice of the β-galactosidase substrate.

The basic protocol for an assay in 96 or 384 well plates is as follows:

A yeast strain containing the *Escherichia coli*

lacZ gene under the control of the yeast Gal4 promoter is grown overnight (with shaking at 230-270 rpm) then diluted with YPD medium to an OD600 of 0.2. Diluted cultures are grown for an additional 3-5 hr until mid-log phase. Yeast cells are then transferred to either 96- or 384-well plates (100 µl/well or 25 µl/well, respectively). Alternatively, cells can be cultured in the microtiter plates, eliminating the need for a pipetting step.

The yeast cells are then either lysed by freeze and thaw method (liquid $N_2$ to freeze, 37°C water bath to thaw) or by use of a Lysis buffer (e.g.: 1% Lithium dodecyl sulphate, 100 mM EDTA and 10 mM Tris-HCl pH 8.0). Non-lysed cells also give a signal, although the variability is increased if the cells are not lysed. Yeast cells can also be permeabilized with various reagents such as isopropanol (15 %).

The substrate sensitivity must be optimised for efficient detection in a screening process. Fluorescein di galactoside (FDG) is a typical low cost fluorescent reagent for the detection of β-galactosidase; it can be used for screening, although autofluorescent compounds can induce a non-desirable background leading to false positives. Alternative substrates are available that become luminescent upon β-galactosidase cleavage, thereby eliminating background problems. An example of such a substrate Galacton-*Star*® from Tropix. Typically about 1µM substrate is added and the plates are incubated at room temperature for 60 minutes. Fluorescence (for FDG) is then measured at 530 nm. It is typically possible to detect as low as 100 cells per well.

As an alternative to the use of β-galactosidase, secreted alkaline phosphatase can be used as a reporter gene. The use of secreted alkaline phosphatase gives equivalent sensitivity to β-galactosidase with the advantage that there is no need

to lyse the cells. Fluorescent substrates for alkaline phosphatase are available commercially from Sigma-Aldrich (Bornem, Belgium) or Molecular Probes (Eugene, OR, USA).

The test compound can be added at various stages of the above procedure. Generally, the compound is added on the plates onto which the yeast are plated. However, the compound can also be added during the second incubation in order to overcome toxicity ptoblems. As a control, it is important to check whether the compound slows down the growth of the yeast. This can be done using turbidity measurements.

<u>Example 6</u>

<u>Detection of *in vivo* protein-protein interactions using fluorescence energy transfer (FRET).</u>

An *in vivo* FRET assay can be conveniently performed using two different mutants of GFP which absorb and emit light at different wavelengths and which have suitably overlapping emission/absorption spectra, such as EGFP (enhanced green fluorescent protein) and EBFP (enhanced blue fluorescent protein). When two such variant GFPs are brought into close proximity, within a few nanometers distance, fluorescence energy transfer (FRET) can be detected. Such transfer is characterized by a reduction of fluorescence intensity of the donor fluorophore (EBFP) and re-emission of fluorescence at the acceptor fluorophore (EGFP) wavelengths. Therefore if each fluorophore is fused to a protein domain known to bind to the other the protein-protein interaction can be monitored *in vivo* using FRET.

In a typical example, the APC binding domain of UNC-5 cloned in fusion with EBFP, whereas APC is cloned in fusion with EGFP in expression vectors suitable for use in the chosen host cell line or organism. When both fusion proteins are expressed in

a cell line or in *C. elegans* it is possible to monitor and quantify their *in vivo* interaction by irradiating the cells/worms with light at 488nm. When the donor and acceptor fluorophore are brought into close proximity by binding of the two fusion proteins fluorescent energy transfer results in a measurable decreased in fluorescence from the fluorescence donor at a wavelength within the emission spectrum of the donor. In simple terms, what is measured is a quenching phenomenon since light emitted by the donor fluorophore is trapped by the acceptor fluorophore. NB- The experiment could also be performed by measuring fluorescence from the acceptor fluorophore but this is often less sensitive.

Plasmid vectors containing both EGFP and EBFP are commercially available from Clontech (Palo Alto, California, USA). Information on the use of these vectors is also supplied by the manufacturer.

Example 7

Genetic and complementation screens in yeast.

UNC-5 expression in yeast cells results in a lethal phenotype, mainly because of the expression of a death domain. This observation was most clearly seen in the experiments with *C. elegans* UNC-5. Accordingly, assays can be developed to screen for compounds, interacting proteins and suppressors which alter the activity of UNC-5, particularly the activity of the death domain of UNC-5. These assays are analogous to those described by Xu and Reed (Mol. Cell 1998, 1:337-46).

(A) Compound screens.

Yeast cells are transfected with a plasmid encoding the *C. elegans* or human unc-5 (including the death domain), such as the vectors described in the

yeast two hybrid experiments. The transfected yeast cells are then placed in the wells of micotiter plates, and are exposed to the compounds under test. Compounds which reduce or inhibit the lethal phenotype of the yeast cells transfected with unc-5 are scored as hits. Such compounds will typically suppress the unc-5 lethal phenotype by interacting with UNC-5 itself, or with UNC-5 interacting proteins, or with proteins in the UNC-5 pathway, or with proteins in parallel pathways. The selected compounds can be used in the development of pharmaceutical preparations.

(B) Suppressor screens.

Yeast cells are transfected with a plasmid encoding the *C. elegans* or human unc-5 (including the death domain), such as the vectors described in the yeast two hybrid experiments. Furthermore, the yeast cells are transfected with a library expressing *C. elegans* or human cDNA, such as the libraries described in the yeast two hybrid experiments. The transfected yeast cells are placed in the wells of micotiter plates, and allowed to grow further. This allows selection cDNAs, and hence genes and proteins, that reduce or inhibit the lethal phenotype of the yeast cells transfected with the death domain of unc-5. Such proteins will interact with UNC-5, or with UNC-5 interacting proteins, or with proteins in the UNC-5 pathway, or with proteins in parallel pathways to cause suppression of the unc-5 lethal phenotype. The selected cDNAs genes and proteins can be used in the development of pharmaceutical preparations or in the development of assays to select for compounds that enhance their function or expression.

Example 8

Cloning of a *C. elegans* gene starting from a *C.*

*elegans* insert.

If a fragment of a given gene or cDNA is known then further fragments of the corresponding full length gene and/or cDNA can be constructed can often be found using *in silico* techniques such as AceDB (see http:\\www.sanger.ac.uk), or searching of the EST database. The full cDNA can be cloned using standard technology such as 5'/3' RACE or SL1/2 RT-PCR on worm total RNA and colony hybridization. An analogous strategy is followed to clone a full length gene and/or cDNA for vertebrate and hence Human DNA.

Example 9

Cloning of *C. elegans* gene starting from a human insert.

A full length *C. elegans* gene can be cloned starting from a human sequence. Using *in silico* techniques, a homologue or an EST can be found. Standard molecular biology techniques can then be used to clone the full length *C. elegans* gene. If no homologous sequence can be found by simple database searching, it may be necessary to perform species hopping. An analogous strategy is followed to clone a full length gene and/or cDNA for vertebrate and hence Human DNA, starting from a *C. elegans* DNA sequence.

**SEQUENCE LISTING**

SEQ ID NO: 1 nucleotide sequence of a part of the human unc-5Cb cDNA which encodes the intracellular region of the protein.

SEQ ID NO: 2 amino acid sequence of the intracellular part of the human unc-5Cb protein encoded by the nucleotide sequence shown as SEQ ID NO: 1.

SEQ ID NO: 3 nucleotide sequence of a part of the human unc-5Cc cDNA which encodes the intracellular region of the protein.

SEQ ID NO: 4 amino acid sequence of the intracellular part of the human unc-5Cc protein encoded by the nucleotide sequence shown as SEQ ID NO: 3.

SEQ ID NO: 5 nucleotide sequence of a part of the human unc-5C8 cDNA which encodes the intracellular region of the protein.

SEQ ID NO: 6 amino acid sequence of the intracellular part of the human unc-5C8 protein encoded by the nucleotide sequence shown as SEQ ID NO: 5.

SEQ ID NO: 7 nucleotide sequence of the fragment of the human unc-5H1 cDNA cloned by PCR in Example 2.

SEQ ID NO: 8 predicted amino acid sequence for the human unc-5H1 protein, translation in frame 1.

SEQ ID NO: 9 predicted amino acid sequence for the human unc-5H1 protein, translation in frame 2.

SEQ ID NO: 10 predicted amino acid sequence for the human unc-5H1 protein, translation in frame 3.

SEQ ID NO: 11 nucleotide sequence of the *C. elegans* spectrin β-chain/Fodrin cDNA.

SEQ ID NO: 12 amino acid sequence of the *C. elegans* spectrin β-chain/Fodrin protein.

SEQ ID NO: 13 nucleotide sequence of the fragment of the *C. elegans* spectrin β-chain/Fodrin cDNA cloned in pC1025.

SEQ ID NO: 14 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 13.

SEQ ID NO: 15 nucleotide sequence of the *C. elegans* APR-1 cDNA.

SEQ ID NO: 16 amino acid sequence of the *C. elegans* APR-1 protein.

SEQ ID NO: 17 nucleotide sequence of a fragment of the *C. elegans* APR-1 cDNA cloned in pC1028.

SEQ ID NO: 18 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 17.

SEQ ID NO: 19 nucleotide sequence of the *C. elegans*

unc-14 cDNA.

SEQ ID NO: 20 amino acid sequence of the *C. elegans* unc-14 protein.

SEQ ID NO: 21 nucleotide sequence of the fragment of the *C. elegans* unc-14 cDNA cloned in pC1034.

SEQ ID NO: 22 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 21.

SEQ ID NO: 23 nucleotide sequence of the *C. elegans* F11A10.1 cDNA.

SEQ ID NO: 24 amino acid sequence of the *C. elegans* F11A10.1 protein.

SEQ ID NO: 25 nucleotide sequence of the fragment of the *C. elegans* F11A10.1 cDNA cloned in pGC1021.

SEQ ID NO: 26 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 25.

SEQ ID NO: 27 nucleotide sequence of the *C.elegans* C15E6.1 cDNA.

SEQ ID NO: 28 amino acid sequence of the *C.elegans* C15E6.1 protein.

SEQ ID NO: 29 nucleotide sequence of the fragment of the *C.elegans* C15E6.1 cDNA cloned in pGC1026.

| | |
|---|---|
| SEQ ID NO: 30 | amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 29. |
| SEQ ID NO: 31 | nucleotide sequence of the *C. elegans* D1081.7 cDNA. |
| SEQ ID NO: 32 | amino acid sequence of the *C. elegans* D1081.7 protein. |
| SEQ ID NO: 33 | nucleotide sequence of the fragment of the *C. elegans* 1081.7 cDNA cloned in pGC1027. |
| SEQ ID NO: 34 | amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 33. |
| SEQ ID NO: 35 | nucleotide sequence of the *C. elegans* B0238.9 cDNA (seu-2). |
| SEQ ID NO: 36 | amino acid sequence of the *C. elegans* B0238.9 protein (seu-2). |
| SEQ ID NO: 37 | nucleotide sequence of the fragment of the *C. elegans* B0238.9 cDNA cloned in pGC1023. |
| SEQ ID NO: 38 | amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 37. |
| SEQ ID NO: 39 | nucleotide sequence of the *C. elegans* ZC404.8 cDNA. |
| SEQ ID NO: 40 | amino acid sequence of the *C. elegans* ZC404.8 protein. |

SEQ ID NO: 41 nucleotide sequence of the *C. elegans* ZC404.8 cDNA cloned in pGC1033.

SEQ ID NO: 42 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 41.

SEQ ID NO: 43 nucleotide sequence of the fragment of the *C. elegans* yk17a3 cDNA cloned in pGC1023.

SEQ ID NO: 44 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 43.

SEQ ID NO: 45 nucleotide sequence of the *C. elegans* F41H10.3 cDNA.

SEQ ID NO: 46 amino acid sequence of the *C. elegans* F41H10.3 protein.

SEQ ID NO: 47 nucleotide sequence of the fragment of the *C. elegans* F41H10.3 cDNA cloned in pGC1020.

SEQ ID NO: 48 amino acid sequence of the polypeptide encoded by the cDNA fragment shown as SEQ ID NO: 47.

SEQ ID NO: 49 nucleotide sequence of the human i-beta-1,3-N-acetylaminyltransferase cDNA.

SEQ ID NO: 50 amino acid sequence of the human i-beta-1,3-N-acetylaminyltransferase protein.

SEQ ID NO: 51 partial nucleotide sequence for the fragment of the human i-beta-1,3-N-acetylaminyltransferase cDNA cloned in pYMP5 (forward primer, coding strand).

SEQ ID NO: 52 partial nucleotide sequence for the fragment of the human i-beta-1,3-N-acetylaminyltransferase cDNA cloned in pYMP5 (reverse primer, non-coding strand)

SEQ ID NO: 53 partial amino acid sequence for the polypeptide encoded by the fragment of the i-beta-1,3-N-acetylaminyltransferase cDNA cloned in pYMP5.

SEQ ID NO: 54 partial nucleotide sequence for the human cDNA fragment cloned in pYMP6 (forward primer, coding strand).

SEQ ID NO: 55 partial nucleotide sequence for the human cDNA fragment cloned in pYMP6 (reverse primer, non-coding strand).

SEQ ID NO: 56 partial nucleotide sequence for the human cDNA fragment cloned in pYMP17 (forward primer, coding strand).

SEQ ID NO: 57 partial nucleotide sequence for the human cDNA fragment cloned in pYMP17 (reverse primer, non-coding strand).

SEQ ID NO: 58 nucleotide sequence of the human alpha-2-macroglobulin cDNA.

SEQ ID NO: 59 amino acid sequence of the human alpha-

2-macroglobulin protein.

SEQ ID NO: 60 partial nucleotide sequence for the fragment of the human alpha-2-macroglobulin cDNA cloned in pYMP30 (reverse primer, non-coding strand).

SEQ ID NO: 61 partial nucleotide sequence of the fragment of human cDNA cloned in pYMP11 (forward primer, coding strand).

SEQ ID NO: 62 partial nucleotide sequence of the fragment of human cDNA cloned in pYMP11 (reverse primer, non-coding strand).

SEQ ID NO: 63 partial nucleotide sequence of the fragment of human cDNA cloned in pYMP12 (forward primer, coding strand).

SEQ ID NO: 64 partial nucleotide sequence of the fragment of human cDNA cloned in pYMP12 (reverse primer, non-coding strand).

SEQ ID NO: 65 amino acid sequence of the mouse APC-2 cDNA.

SEQ ID NO: 66 nucleotide sequence of a *C. elegans* I-beta-1,3-N-acetylaminyltransferase cDNA (F22F7.6).

SEQ ID NO: 67 amino acid sequence of a *C. elegans* I-beta-1,3-N-acetylaminyltransferase protein (F22F7.6).

SEQ ID NO: 68 nucleotide sequence of the *C. elegans* alpha-2-macroglobulin cDNA ZK337.1a.

SEQ ID NO: 69 nucleotide sequence of the *C. elegans* alpha-2-macroglobulin cDNA ZK337.1b

SEQ ID NO: 70 amino acid sequence of the *C. elegans* alpha-2-macroglobulin protein ZK337.1a.

SEQ ID NO: 71 amino acid sequence of the *C. elegans* alpha-2-macroglobulin protein ZK337.1b.

SEQ ID NO: 72 cDNA sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue C18C1.3.

SEQ ID NO: 73 amino acid sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue C18C1.3.

SEQ ID NO: 74 cDNA sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue K09C8.4.

SEQ ID NO: 75 amino acid sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue K09C8.4.

SEQ ID NO: 76 amino acid sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue F21H7.10.

SEQ ID NO: 77 cDNA sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue C54C8.2.

SEQ ID NO: 78 amino acid sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue C54C8.2.

SEQ ID NO: 79 cDNA sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue F56H6.6.

SEQ ID NO: 80 amino acid sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue F56H6.6.

SEQ ID NO: 81 cDNA sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue T15D6.4.

SEQ ID NO: 82 amino acid sequence for the *C. elegans* I-beta-1,3-N-acetylaminyltransferase homologue T15D6.4.

SEQ ID NO: 83 amino acid sequence of the extracellular part of the *C. elegans* unc-5 protein.

SEQ ID NO: 84 amino acid sequence of the transmembrane region of the *C. elegans* unc-5 protein.

SEQ ID NO: 85 amino acid sequence of the membrane proximal part of the *C. elegans* unc-5 protein.

SEQ ID NO: 86 amino acid sequence of the zonula occludens part of the *C. elegans* unc-5 protein.

SEQ ID NO: 87 amino acid sequence of a part of the *C. elegans* unc-5 protein of unknown function.

SEQ ID NO: 88 amino acid sequence of the death domain

of the *C. elegans* unc-5 protein.

SEQ ID NO: 89 amino acid sequence of the human HS1 protein.

SEQ ID NO: 90 amino acid sequence of the human UNC5C protein.

SEQ ID NO: 91 complete nucleotide sequence of plasmid pGC1037.

SEQ ID NO: 92 complete nucleotide sequence of plasmid pGC1003.

SEQ ID NO: 93 amino acid sequence of *C. elegans* unc-40.

SEQ ID NO: 94 nucleotide sequence of *C. elegans* unc-40.

SEQ ID NO: 95 amino acid sequence of human unc-40.

SEQ ID NO: 96 nucleotide sequence of human unc-40.

**ACCESSION NUMBERS:**

Human beta-fodrin cDNA-GenBank S65762

Human beta-fodrin protein-swissprot Q01082

Human APC-1 cDNA-GenBank M74088

Human APC-1 protein-swissprot P25054

Human unc-14 cDNA (KIAA0375)-GenBank AB002373

Human unc-14 protein (KIAA0375)-BAA20830

Human yk17a3 cDNA (KIAA0187)-GenBank D80009

Human yk17a3 protein (KIAA0187)-SPTREMBL:Q14692

TABLE 1: Schematic representation of dimerisations of C. elegans unc-5, using constructions in pAS2 and pGAD424

| | pGAD424 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| pAS2 | full length unc-5 (1016) | Dd (1008) | MPP (1009) | MPP + ZO-1 (1010) | MPP + ZO-1 + UP (1011) | UP (1013) | ZO-1 (1012) | empty pGAD424 |
| full length unc-5 (1006) | nd | nd | nd | nd | nd | nd | nd | not blue |
| UP + DD (1000) auto-activation | nd | **blue** | nd | nd | nd | nd | nd | **blue** |
| MPP (1001) | nd | nd | nd | nd | nd | nd | nd | nd |
| MPP + ZO-1 (1002) | not blue | nd | nd | nd | nd | nd | nd | nd |
| MPP + ZO-1 + UP (1003) | not blue | not blue | not blue | nd | **blue** | not blue | not blue | not blue |
| ZO-1 (1007) | nd | nd | nd | nd | nd | nd | not blue | nd |
| UP (1004) | nd | nnd | nd | nd | nd | not blue | nd | nd |
| ZO-1 + UP (1005) | not blue | nd | nd | nd | nd | nd | **blue** | nd |
| empty pAS2 | not blue | nd | nd | nd | nd | nd | nd | nd |

**Claims:**

1. A protein comprising the sequence of amino acids set forth in SEQ ID NO: 2 or a sequence of amino acids which differs from that set forth in SEQ ID NO: 2 only in conservative amino acid changes.

2. A nucleic acid comprising a sequence of nucleotides which encodes the protein claimed in claim 1.

3. A nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 1 or a fragment thereof.

4. An expression vector comprising the nucleic acid of claim 2 or claim 3.

5. A host cell or organism transformed or transfected with the expression vector of claim 4.

6. An antibody which is capable of specifically binding to the protein claimed in claim 1 or an epitope thereof.

7. A protein comprising the sequence of amino acids set forth in SEQ ID NO: 4 or a sequence of amino acids which differs from that set forth in SEQ ID NO: 4 only in conservative amino acid changes.

8. A nucleic acid comprising a sequence of nucleotides which encodes the protein claimed in claim 7.

9. A nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 3 or a fragment thereof.

10. An expression vector comprising the nucleic acid of claim 8 or claim 9.

11. A host cell or organism transformed or transfected with the expression vector of claim 10.

12. An antibody which is capable of specifically binding to the protein claimed in claim 7 or an epitope thereof.

13. A protein comprising the sequence of amino acids set forth in SEQ ID NO: 6 or a sequence of amino acids which differs from that set forth in SEQ ID NO: 6 only in conservative amino acid changes.

14. A nucleic acid comprising a sequence of nucleotides which encodes the protein claimed in claim 13.

15. A nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 5 or a fragment thereof.

16. An expression vector comprising the nucleic acid of claim 13 or claim 14.

17. A host cell or organism transformed or transfected with the expression vector of claim 16.

18. An antibody which is capable of specifically binding to the protein claimed in claim 13 or an epitope thereof.

19. A method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which

method comprises:

providing a host cell containing a DNA construct comprising a reporter gene operatively linked to a promoter regulated by a transcription factor having a DNA binding domain and an activating domain;

expressing in said host cell a first hybrid DNA sequence encoding a first fusion protein comprising an UNC-5 protein or a fragment thereof fused in-frame to either the DNA binding domain or the activating domain of the said transcription factor;

expressing in said host cell a second hybrid DNA sequence encoding a second fusion protein comprising an interacting protein or a fragment thereof fused in-frame to either the DNA binding domain or the activating domain of the said transcription factor, such that when the first fusion protein comprises the activation domain of the said transcription factor the second fusion protein comprises the DNA binding domain of the said transcription factor and when the first fusion protein comprises the DNA binding domain of the transcription factor the second fusion protein comprises the activation domain;

contacting the host cell with a sample of the compound under test; and

detecting any binding of the UNC-5 protein or fragment thereof to the interacting protein or fragment thereof by detecting the production of any reporter gene product in the said host cell.

20. A method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

providing a transgenic cell or organism expressing a first fusion protein comprising an UNC-5 protein or a fragment thereof fused in-frame to a first genetically encoded fluorophore and a second fusion protein comprising an interacting protein or a fragment thereof fused in-frame to a second genetically encoded fluorophore, the first and second fluorophores being characterised in that the emission spectrum of one of the fluorophores overlaps with the absorption spectrum of the other fluorophore;

measuring the amount of fluorescence emitted from the fluorophore having an emission spectrum which overlaps with the absorption spectrum of the other fluorophore;

exposing the transgenic cell or organism to a compound under test; and

detecting any change in the amount of fluorescence emitted fluorescence emitted from the fluorophore having an emission spectrum which overlaps with the absorption spectrum of the other fluorophore.

21. A method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

providing a first reaction component comprising a first protein linked to a solid support containing a scintillant and a second reaction component comprising a second protein which has been radioactively labelled, wherein the first and second proteins are an UNC-5 protein or a fragment thereof and an interacting protein or a fragment thereof;

bringing the first and second reaction

components into contact in an aqueous solution in the presence of a compound under test; and

detecting binding of the first protein to the second protein by detecting light emission from the scintillant.

22. A method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

coating the wells of a microtiter plate with UNC-5 protein or a fragment thereof;

contacting the UNC-5 protein or fragment thereof with an aqueous solution comprising an interacting protein or a fragment thereof, said interacting protein being labelled with a tag which is directly or indirectly detectable, and a compound under test;

washing to remove the compound under test and any unbound tagged interacting protein; and

detecting complexes of UNC-5 or a fragment thereof bound to the interacting protein or a fragment thereof by directly or indirectly detecting the presence of the tag.

23. A method of identifying compounds which are capable of inhibiting or enhancing the binding of an UNC-5 protein to an interacting protein previously identified as binding to the said UNC-5 protein, which method comprises:

exposing a cell or organism expressing UNC-5 and overexpressing nucleic acid encoding an interacting protein to the compound under test; and

screening for reversion of the overexpression phenotype of the cell or organism

to wild-type.

24. A method as claimed in claim 23 wherein the organism is a nematode worm.

25. A method as claimed in claim 24 wherein the nematode worm is *C. elegans*.

26. A method as claimed in claim 23 wherein the cell is a mammalian cell line.

27. A method as claimed in any one of claims 23 to 26 wherein the cell or organism further expresses a reporter gene encoding a reporter protein.

28. A method as claimed in claim 27 wherein the reporter protein is a fluorescent protein or a luminescent protein.

29. A method as claimed in any one of claims 19 to 28 wherein the UNC-5 protein is a *C. elegans* UNC-5 protein.

30. A method as claimed in any one of claims 19 to 28 wherein the UNC-5 protein is a human UNC-5 protein.

31. A method as claimed in claim 30 wherein the human UNC-5 protein is UNC-5C or a protein as claimed in any one of claims 1, 7, 13 or 71.

32. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is a *C. elegans* UNC-5 protein or a fragment thereof.

33. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is *C. elegans*

UNC-40.

34. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is human UNC-40.

35. A method as claimed in claim 34 wherein the UNC-40 protein comprises the sequence of amino acids set forth in SEQ ID NO: 95.

36. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is a *C. elegans* spectrin β-chain/fodrin protein.

37. A method as claimed in claim 36 wherein the spectrin β-chain/fodrin protein comprises the sequence of amino acids set forth in SEQ ID NO: 12.

38. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is *C. elegans* APR-1.

39. A method as claimed in claim 38 wherein the *C. elegans* APR-1 protein comprises the sequence of amino acids set forth in SEQ ID NO: 16.

40. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is *C. elegans* UNC-14.

41. A method as claimed in claim 40 wherein the *C. elegans* UNC-14 protein comprises the sequence of amino acids set forth in SEQ ID NO: 20.

42. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of amino acids set forth in SEQ ID NO: 24.

43. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of amino acids set forth in SEQ ID NO: 28.

44. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of nucleotides set forth in SEQ ID NO: 32.

45. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of amino acids set forth in SEQ ID NO: 36.

46. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of amino acids set forth in SEQ ID NO: 40.

47. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of amino acids set forth in SEQ ID NO: 44.

48. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises the sequence of amino acids set forth in SEQ ID NO: 46.

49. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is a human UNC-5 protein.

50. A method as claimed in claim 49 wherein the human UNC-5 protein is UNC-5C or a protein as claimed in any one of claims 1, 7, 13 or 71.

51. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is human i-beta-1,3-N-acetylaminyltransferase.

52. A method as claimed in claim 51 wherein the

human i-beta-1,3-N-acetylaminyltransferase comprises the sequence of amino acids set forth in SEQ ID NO: 50.

53. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises a protein encoded by the nucleic acid of claim 72 or claim 73.

54. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises a protein encoded by the nucleic acid of claim 74 or claim 75.

55. A method as claimed in any one of claims 29 to 31 wherein the interacting protein is human alpha-2 macroglobulin.

56. A method as claimed in claim 55 wherein the alpha-2 macroglobulin comprises the sequence of amino acids set forth in SEQ ID NO: 59.

57. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises a protein encoded by the nucleic acid of claim 76 or claim 77.

58. A method as claimed in any one of claims 29 to 31 wherein the interacting protein comprises a protein encoded by the nucleic acid of claim 78 or claim 79.

59. A method of identifying compounds which reduce or inhibit the lethal phenotype associated with the expression of the UNC-5 death domain in yeast, which method comprises:

exposing a yeast cell containing an

expression vector comprising nucleic acid encoding an UNC-5 protein or a fragment thereof comprising the death domain to a compound under test;

allowing the yeast cells to grow in the presence of the compound; and

screening for a reduction or inhibition of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast.

60. A method as claimed in claim 59 wherein the UNC-5 protein is a *C. elegans* UNC-5 protein.

61. A method as claimed in claim 59 wherein the UNC-5 protein is a human UNC-5 protein.

62. A method as claimed in claim 61 wherein the human UNC-5 protein is a protein as claimed in any one of claims 1, 7 or 13 or 71.

63. A method of identifying suppressers of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast, which method comprises:

transfecting yeast cells containing an expression vector comprising nucleic acid encoding an UNC-5 protein or a fragment thereof comprising the death domain with a cDNA library cloned in a yeast expression vector;

allowing the transfected yeast cells to grow for one or more cell divisions; and

screening for reduction or inhibition of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast.

64. A method as claimed in claim 63, which method further comprises the steps of:

identifying a transfected yeast cell

exhibiting a reduction or inhibition of the lethal phenotype associated with the expression of the UNC-5 death domain in yeast; and

isolating the cDNA clone(s) present in the transfected yeast cell which is/are responsible for conferring reduction or inhibition of the lethal phenotype.

65. A method as claimed in claim 63 or claim 64 wherein the UNC-5 protein is a *C. elegans* UNC-5 protein.

66. A method as claimed in claim 63 or claim 64 wherein the UNC-5 protein is a human UNC-5 protein.

67. A method as claimed in claim 66 wherein the human UNC-5 protein is a protein as claimed in any one of claims 1, 7, 13 or 71.

68. A method as claimed in claim 65 wherein the cDNA library is a *C. elegans* cDNA library.

69. A method as claimed in claim 66 or claim 67 wherein the cDNA library is a human cDNA library.

70. A nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 7.

71. A protein comprising a sequence of amino acids encoded by the nucleic acid molecule of claim 8.

72. A nucleic acid which is obtainable by restriction enzyme digestion of the plasmid pYMP17 with the restriction enzymes EcoRI and XhoI.

73. A nucleic acid as claimed in claim 72 which comprises the sequence of nucleotides set forth in SEQ

ID NO: 56 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 57.

74. A nucleic acid which is obtainable by restriction enzyme digestion of the plasmid pYMP6 with the restriction enzymes EcoRI and XhoI.

75. A nucleic acid as claimed in claim 74 which comprises the sequence of nucleotides set forth in SEQ ID NO: 54 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 55.

76. A nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 61 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 62.

77. A nucleic acid as claimed in claim 76 which is obtainable by restriction enzyme digestion of the plasmid pYMP11 with the restriction enzymes EcoRI and XhoI.

78. A nucleic acid comprising the sequence of nucleotides set forth in SEQ ID NO: 63 and a sequence of nucleotides complementary to the sequence of nucleotides set forth in SEQ ID NO: 64.

79. A nucleic acid as claimed in claim 78 which is obtainable by restriction enzyme digestion of the plasmid pYMP12 with the restriction enzymes EcoRI and XhoI.

80. A nucleic acid probe which is capable of hybridizing to the nucleic acid of claim 70 under conditions of high stringency.

81. An oligonucleotide comprising a sequence of 10 or more consecutive nucleotides of the sequence of nucleotides set forth in SEQ ID NO: 7.

82. An antisense nucleic acid which is capable of hybridizing to the sequence of nucleotides set forth in SEQ ID NO: 7 under conditions of high stringency.

83. An expression vector comprising the nucleic acid of claim 70.

84. A host cell or organism transformed or transfected with the expression vector of claim 83.

85. An antibody which is capable of specifically binding to the protein claimed in claim 71.

FIG. 1.

```
Multalin version 5.3.3
Copyright I.N.R.A. France 1989, 1991, 1994, 1996
Published research using this software should cite
Multiple sequence alignment with hierarchical clustering
F. CORPET, 1988, Nucl. Acids Res., 16 22, 10881-10890
Symbol comparison table: blosum62
Gap weight: 12
Gap length weight: 2
Consensus levels: high=90% low=50%
Consensus symbols:
 ! is anyone of IV
 $ is anyone of LM
 % is anyone of FY
 # is anyone of NDQEBZ

 MSF:   1599    Check:   0
 Name: UNC5C           Len:  1599  Check:  410  Weight:  0.76
 Name: UNC5C8          Len:  1599  Check: 1710  Weight:  0.76
 Name: UNC5Cc          Len:  1599  Check: 5512  Weight:  1.12
 Name: UNC5Cd(UNC5Cb)  Len:  1599  Check: 1388  Weight:  1.37
 Name: Consensus       Len:  1599  Check: 7845  Weight:  4.00


           1                                                   50
    UNC5C  TTGTTTGTGT ATCGGAAGAA TCATCGTGAC TTTGAGTCAG ATATTATTGA
   UNC5C8  TTGTTTGTGT ATCGGAAGAA TCATCGTGAC TTTGAGTCAG ATATTATTGA
   UNC5Cc  TTGTTTGTGT ATCGGAAGAA TCATCGTGAC TTTGAGTCAG ATATTATTGA
   UNC5Cd  TTGTTTGTGT ATCGGAAGAA TCATCGTGAC TTTGAGTCAG ATATTATTGA
Consensus  TTGTTTGTGT ATCGGAAGAA TCATCGTGAC TTTGAGTCAG ATATTATTGA

           51                                                 100
    UNC5C  CTCTTCGGCA CTCAATGGGG GCTTTCAGCC TGTGAACATC AAGGCAGCAA
   UNC5C8  CTCTTCGGCA CTCAATGGGG GCTTTCAGCC TGTGAACATC AAGGCAGCAA
   UNC5Cc  CTCTTCGGCA CTCAATGGGG GCTTTCAGCC TGTGAACATC AAGGCAGCAA
   UNC5Cd  CTCTTCGGCA CTCAATGGGG GCTTTCAGCC TGTGAACATC AAGGCAGCAA
Consensus  CTCTTCGGCA CTCAATGGGG GCTTTCAGCC TGTGAACATC AAGGCAGCAA

           101                                                150
    UNC5C  GACAAGATCT GCTGGCTGTA CCCCCAGACC TCACGTCAGC TGCAGCCATG
   UNC5C8  GACAAGA--- ---------- --------CC TCACGTCAGC TGCAGCCATG
   UNC5Cc  GACAAGATCT GCTGGCTGTA CCCCCAGACC TCACGTCAGC TGCAGCCATG
   UNC5Cd  GACAAGATCT GCTGGCTGTA CCCCCAGACC TCACGTCAGC TGCAGCCATG
Consensus  GACAAGAtct gctggctgta cccccagaCC TCACGTCAGC TGCAGCCATG

           151                                                200
    UNC5C  TACAGAGGAC CTGTCTATGC CCTGCATGAC GTCTCAGACA AAATCCCAAT
   UNC5C8  TACAGAGGAC CTGTCTATGC CCTGCATGAC GTCTCAGACA AAATCCCAAT
   UNC5Cc  TACAGAGGAC CTGTCTATGC CCTGCATGAC GTCTCAGACA AAATCCCAAT
   UNC5Cd  TACAGAGGAC CTGTCTATGC CCTGCATGAC GTCTCAGACA AAATCCCAAT
Consensus  TACAGAGGAC CTGTCTATGC CCTGCATGAC GTCTCAGACA AAATCCCAAT

           201                                                250
    UNC5C  GACCAACTCT CCAATTCTGG ATCCACTGCC CAACCTGAAA ATCAAAGTGT
   UNC5C8  GACCAACTCT CCAATTCTGG ATCCACTGCC CAACCTGAAA ATCAAAGTGT
   UNC5Cc  GACCAACTCT CCAATTCTGG ATCCACTGCC CAACCTGAAA ATCAAAGTGT
   UNC5Cd  GACCAACTCT CCAATTCTGG ATCCACTGCC CAACCTGAAA ATCAAAGTGT
Consensus  GACCAACTCT CCAATTCTGG ATCCACTGCC CAACCTGAAA ATCAAAGTGT

           251                                                300
    UNC5C  ACAACACCTC AGGTGCTGTC TCCCCCCAAG ATGACCTCTC TGAGTTTACG
   UNC5C8  ACAACACCTC AGGTGCTGTC ACCCCCCAAG ATGACCTCTC TGAGTTTACG
   UNC5Cc  ACAACACCTC AGGTGCTGTC ACC------- ---------- ----------
   UNC5Cd  ACAACACCTC AAGTGCTGTC ACCCCCCAAG ATGACCTCTC TGAGTTTACG
Consensus  ACAACACCTC AgGTGCTGTC aCCccccaag atgacctctc tgagtttacg

           301                                                350
    UNC5C  TCCAAGCTGT CCCCTCAGAT GACCCAGTCG TTGTTGGAGA ATGAAGCCCT
   UNC5C8  TCCAAGCTGT CCCCTCAGAT GACCCAGTCG TTGTTGGAGA ATGAAGCCCT
   UNC5Cc  ---------- ---------- ---------- ---------- ----------
   UNC5Cd  TCCAAGCTGT CCCCTCAGAT GACCCAGTCG TTGTTGGAGA ATGAAGCCCT
Consensus  tccaagctgt cccctcagat gacccagtcg ttgttggaga atgaagccct
```

FIG. 1 (CONTINUED 1)

```
           351                                                    400
UNC5C      CAGCCTGAAG AACCAGAGTC TAGCAAGGCA GACTGATCCA TCCTGTACCG
UNC5C8     CAGCCTGAAG AACCAGAGTC TAGCAAGGCA GACTGATCCA TCCTGTACCG
UNC5Cc     ---------- ---------- ---------- ---------- ----------
UNC5Cd     CAGCCTGAAG AACCAGAGTC TAGCAAGGCA GACTGATCCA TCCTGTACCG
Consensus  cagcctgaag aaccagagtc tagcaaggca gactgatcca tcctgtaccg

           401                                                    450
UNC5C      CATTTGGCAG CTTCAACTCG CTGGGAGGTC ACCTTATTGT TCCCAATTCA
UNC5C8     CATTTGGCAG CTTCAACTCG CTGGGAGGTC ACCTTATTGT TCCCAATTCA
UNC5Cc     ---------- ---------- ---------- ----TATTGT TCCCAATTCA
UNC5Cd     CATTTGGCAG CTTCAACTCG CTGGGAGGTC ACCTTATTGT TCCCAATTCA
Consensus  catttggcag cttcaactcg ctgggaggtc acctTATTGT TCCCAATTCA

           451                                                    500
UNC5C      GGAGTCAGCT TGCTGATTCC CGCTGGGGCC ATTCCCCAAG GGAGAGTCTA
UNC5C8     GGAGTCAGCT TGCTGATTCC CGCTGGGGCC ATTCCCCAAG GGAGAGTCTA
UNC5Cc     GGAGTCAGCT TGCTGATTCC CGCTGGGGCC ATTCCCCAAG GGAGAGTCTA
UNC5Cd     GGAGTCAGCT TGCTGATTCC CGCTGGGGCC ATTCCCCAAG GGAGAGTCTA
Consensus  GGAGTCAGCT TGCTGATTCC CGCTGGGGCC ATTCCCCAAG GGAGAGTCTA

           501                                                    550
UNC5C      CGAAATGTAT GTGACTGTAC ACAGGAAAGA AACTATGAGG CCACCCATGG
UNC5C8     CGAAATGTAT GTGACTGTAC ACAGGAAAGA AACTATGAGG CCACCCATGG
UNC5Cc     CGAAATGTAT GTGACTGTAC ACAGGAAAGA AACTATGAGG CCACCCATGG
UNC5Cd     CGAAATGTAT GTGACTGTAC ACAGGAAAGA AACTATGAGG CCACCCATGG
Consensus  CGAAATGTAT GTGACTGTAC ACAGGAAAGA AACTATGAGG CCACCCATGG

           551                                                    600
UNC5C      ATGACTCTCA GACACTTTTG ACCCCTGTGG TGAGCTGTGG GCCCCCAGGA
UNC5C8     ATGACTCTCA GACACTTTTG ACCCCTGTGG TGAGCTGTGG GCCCCCAGGA
UNC5Cc     ATGACTCTCA GACACTTTTG ACCCCTGTGG TGAGCTGTGG GCCCCCAGGA
UNC5Cd     ATGACTCTCA GACACTTTTG ACCCCTGTGG TGAGCTGTGG GCCCCCAGGA
Consensus  ATGACTCTCA GACACTTTTG ACCCCTGTGG TGAGCTGTGG GCCCCCAGGA

           601                                                    650
UNC5C      GCTCTGCTCA CCCGCCCCGT CGTCCTCACT ATGCATCACT GCGCAGACCC
UNC5C8     GCTCTGCTCA CCCGCCCCGC CGTCCTCACT ATGCATCACT GCGCAGACCC
UNC5Cc     GCTCTGCTCA CCCGCCCCGT CGTCCTCACT ATGCATCACT GCGCAGACCC
UNC5Cd     GCTCTGCTCA CCCGCCCCGT CGTCCTCACT ATGCATCACT GCGCAGACCC
Consensus  GCTCTGCTCA CCCGCCCCGt CGTCCTCACT ATGCATCACT GCGCAGACCC

           651                                                    700
UNC5C      CAATACCGAG GACTGGAAAA TACTGCTCAA GAACCAGGCA GCACAGGGAC
UNC5C8     CAATACCGAG GACTGGAAAA TACTGCTCAA GAACCAGGCA GCACAGGGAC
UNC5Cc     CAATACCGAG GACTGGAAAA TACTGCTCAA GAACCAGGCA GCACAGGGAC
UNC5Cd     CAATACCGAG GACTGGAAAA TACTGCTC-- ---------- ----------
Consensus  CAATACCGAG GACTGGAAAA TACTGCTCaa gaaccaggca gcacagggac

           701                                                    750
UNC5C      AGTGGGAGGA TGTGGTGGTG GTCGGGGAGG AAAACTTCAC CACCCCCTGC
UNC5C8     AGTGGGAGGA TGTGGTGGTG GCCGGGGAGG AAAACTTCAC CACCCCCTGC
UNC5Cc     AGTGGGAGGA TGTGGTGGTG GTCGGGGAGG AAAACTTCAC CACCCCCTGC
UNC5Cd     ---------- ---------- ---------- ---------- ----------
Consensus  agtgggagga tgtggtggtg g cggggagg aaaacttcac caccccctgc

           751                                                    800
UNC5C      TACATTAAGC TGGATGCAGA GGCCTGCCAC ATCCTCACAG AGAACCTCAG
UNC5C8     TACATTAAGC TGGATGCAGA GGCCTGCCAC ATCCTCACAG AGAACCTCAG
UNC5Cc     TACATTCAGC TGGATGCAGA GGCCTGCCAC ATCCTCACAG AGAACCTCAG
UNC5Cd     ---------- ---------- ---------- ---------- ----------
Consensus  tacatt agc tggatgcaga ggcctgccac atcctcacag agaacctcag

           801                                                    850
UNC5C      CACCTACGCC CTGGTAGGAC ATTCCACCAC CAAAGCGGCT GCAAAGCGCC
UNC5C8     CACCTACGCC CTGGTAGGAC ATTCCACCAC CAAAGCGGCT GCAAAGCGCC
UNC5Cc     CACCTACGCC CTGGTAGGAC ATTCCACCAC CAAAGCGGCT GCAAAGCGCC
UNC5Cd     ---------- ---------- ---------- ---------- ----------
Consensus  cacctacgcc ctggtaggac attccaccac caaagcggct gcaaagcgcc

           851                                                    900
UNC5C      TCAAGCTGGC CATCTTTGGG CCCCTGTGCT GCTCCTCGCT GGAGTACAGC
UNC5C8     TCAAGCTGGC CATCTTTGGG CCCCTGTGCT GCTCCTCGCT GGAGTACAGC
UNC5Cc     TCAAGCTGGC CATCTTTGGG CCCCTGTGCT GCTCCTCGCT GGAGTACAGC
UNC5Cd     ---------- ---------- ---------- ----CTCGCT GGAGTACAGC
Consensus  tcaagctggc catctttggg cccctgtgct gctcCTCGCT GGAGTACAGC
```

FIG. 1 (CONTINUED 2).

```
           901                                                950
    UNC5C  ATCCGAGTCT ACTGTCTGGA TGACACCCAG GATGCCCTGA AGGAAATTTT
   UNC5C8  ATCCGAGTCT ACTGTCTGGA TGACACCCAG GATGCCCTGA AGGAAATTTT
   UNC5Cc  ATCCGAGTCT ACTGTCTGGA TGACACCCAG GGTGCCCTGA AGGAAATTTT
   UNC5Cd  ATCCGAGTCT ACTGTCTGGA TGACACCCAG GATGCCCTGA AGGAAATTTT
Consensus  ATCCGAGTCT ACTGTCTGGA TGACACCCAG GaTGCCCTGA AGGAAATTTT

           951                                               1000
    UNC5C  ACATCTTGAG AGACAGACGG GAGGACAGCT CCTAGAAGAA CCTAAGGCTC
   UNC5C8  ACATCTTGAG AGAXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXAGGTTT
   UNC5Cc  ACATCTTGAG AGACAGACGG GAGGACAGCT CCTAGAAGAA CCTAAGGCTC
   UNC5Cd  ACATCTTGAG AGACAGACGG GAGGACAGCT CCTAGAAGAA CCTAAGGCTC
Consensus  ACATCTTGAG AGAcagacgg gaggacagct cctagaagaa cctaAGGcTc

           1001                                              1050
    UNC5C  TTCATTTTAA AGGCAGCACC CACAACCTGC GCCTGTCAAT TCACGATATC
   UNC5C8  TTCATTT-AA AGCANGCANC CNNCAAATGN GCCTGTCAAT TCNCGATATG
   UNC5Cc  TTCATTTTAA AGGCAGCACC CACAACCTGC GCCTGTCAAT TCACGATATC
   UNC5Cd  TTCATTTTAA AGGCAGCACC CACAACCTGC GCCTGTCAAT TCACGATATC
Consensus  TTCATTTtAA AGgcaGCAcC CacaAccTGc GCCTGTCAAT TCaCGATATc

           1051                                              1100
    UNC5C  GCCCATTCCC TCTGGAAGAG CAAATTGCTG GCTAAATATC AGGAAATTCC
   UNC5C8  GCCCATTCCC TCTGAAAGAG CAAATTGCTG GCTAAATATC AGGAAATTCC
   UNC5Cc  GCCCGTTCCC TCTGGAAGAG CAAATTGCTG GCTAAATATC AGGAAATTCC
   UNC5Cd  GCCCATTCCC TCTGGAAGAG CAAATTGCTG GCTAAATATC AGGAAATTCC
Consensus  GCCCaTTCCC TCTGgAAGAG CAAATTGCTG GCTAAATATC AGGAAATTCC

           1101                                              1150
    UNC5C  ATTTTACCAT GTTTGGAGTG GATCTCAAAG AAACCTGCAC TGCACCTTCA
   UNC5C8  ATTTTACCAT GTTTGGAGTG GATCTCAAAG AANCCTGCAC TGCACNTTCA
   UNC5Cc  ATTTTACCAT GTTTGGAGTG GATCTCAAAG AAACCTGCAC TGCACCTTCA
   UNC5Cd  ATTTTACCAT GTTTGGAGTG GATCTCAAAG AAACCTGCAC TGCACCTTCA
Consensus  ATTTTACCAT GTTTGGAGTG GATCTCAAAG AAaCCTGCAC TGCACcTTCA

           1151                                              1200
    UNC5C  CTCTGGAAAG ATTTAGCCTG AACACAGTGG AGCTGGTTTG CAAACTCTGT
   UNC5C8  CTCTGGAAAG ATTTAGCCTG AACACAGTGG AGCTGGTTTG CAAACTCTGT
   UNC5Cc  CTCTGGAAAG ATTTAGCCTG AACACAGTGG AGCTGGTTTG CAAACTCTGT
   UNC5Cd  CTCTGGAAAG ATTTAGCCTG AACACAGTGG AGCTGGTTTG CAAACTCTGT
Consensus  CTCTGGAAAG ATTTAGCCTG AACACAGTGG AGCTGGTTTG CAAACTCTGT

           1201                                              1250
    UNC5C  GTGCGGCAGG TGGAAGGAGA AGGGCAGATC TTCCAGCTCA ACTGCACCGT
   UNC5C8  GT-CGGCAGG TGGAAGGAGA AGG-CAGATC TTCCAGCTCA ACTGCACCGT
   UNC5Cc  GTGCGGCAGG TGGAAGGAGA AGGGCAGATC TTCCAGCTCA ACTGCACCGT
   UNC5Cd  GTGCGGCAGG TGGAAGGAGA AGGGCAGATC TTCCAGCTCA ACTGCACCGT
Consensus  GTgCGGCAGG TGGAAGGAGA AGGgCAGATC TTCCAGCTCA ACTGCACCGT

           1251                                              1300
    UNC5C  GTCAGAGGAA CCTACTGGCA TCGATTTGCC GCTGCTGGAT CCTGCGAACA
   UNC5C8  GTCAGAGGAA CCTACTGGCA TCGATTTGCC GCTGCTGGAT CCTGCGAACA
   UNC5Cc  GTCAGAGGAA CCTACTGGCA TCGATTTGCC GCTGCTGGAT CCTGCGAACA
   UNC5Cd  GTCAGAGGAA CCTACTGGCA TCGATTTGCC GCTGCTGGAT CCTGCGAACA
Consensus  GTCAGAGGAA CCTACTGGCA TCGATTTGCC GCTGCTGGAT CCTGCGAACA

           1301                                              1350
    UNC5C  CCATCACCAC GGTCACGGGG CCCAGTGCTT TCAGCATCCC TCTCCCTATC
   UNC5C8  CCATCACCAC GGTCACGGGG CCCAGTGCTT TCAGCATCCC TCTCCCTATC
   UNC5Cc  CCATCACCAC GGTCACGGGG CCCAGTGCTT TCAGCATCCC TCTCCCTATC
   UNC5Cd  CCATCACCAC GGTCACGGGG CCCAGTGCTT TCAGCATCCC TCTCCCTATC
Consensus  CCATCACCAC GGTCACGGGG CCCAGTGCTT TCAGCATCCC TCTCCCTATC

           1351                                              1400
    UNC5C  CGGCAGAAGC TCTGTAGCAG CCTGGATGCC CCCCAGACGA GAGGCCATGA
   UNC5C8  CGGCAGAAGC TCTGTAGCAG CCTGGATGCC CCCCAGACGA GAGGCCATGA
   UNC5Cc  CGGCAGAAGC TCTGTAGCAG CCTGGATGCC CCCCAGACGA GAGGCCATGA
   UNC5Cd  CGGCAGAAGC TCTGTAGCAG CCTGGATGCC CCCCAGACGA GAGGCCATGA
Consensus  CGGCAGAAGC TCTGTAGCAG CCTGGATGCC CCCCAGACGA GAGGCCATGA

           1401                                              1450
    UNC5C  CTGGAGGATG CTGGCCCATA AGCTGAACCT GGACAGGTAC TTGAATTACT
   UNC5C8  CTGGAGGATG CTGGCCCATA AGCTGAACCT GGACGGGTAC TTGAATTACT
   UNC5Cc  CTGGAGGATG CTGGCCCATA AGCTGAACCT GGACAGGTAC TTGAATTACT
   UNC5Cd  CTGGAGGATG CTGGCCCATA AGCTGAACCT GGACAGGTAC TTGAATTACT
Consensus  CTGGAGGATG CTGGCCCATA AGCTGAACCT GGACaGGTAC TTGAATTACT
```

FIG. 1 (CONTINUED 3).

| | 1451 | | | | 1500 |
|---|---|---|---|---|---|
| UNC5C | TTGCCACCAA | ATCCAGCCCA | ACTGGCGTAA | TCCTGGATCT | TTGGGAAGCA |
| UNC5C8 | TTGCCACCAA | ATCCAGCCCA | ACTGGCGTAA | TCCTGGATCT | TTGGGAAGCA |
| UNC5Cc | TTGCCACCAA | ATCCAGCCCA | ACTGGCGTAA | TCCTGGATCT | TTGGGAAGCA |
| UNC5Cd | TTGCCACCAA | ATCCAGCCCA | ACTGGCGTAA | TCCTGGATCT | TTGGGAAGCA |
| Consensus | TTGCCACCAA | ATCCAGCCCA | ACTGGCGTAA | TCCTGGATCT | TTGGGAAGCA |

| | 1501 | | | | 1550 |
|---|---|---|---|---|---|
| UNC5C | CAGAACTTCC | CAGATGGAAA | CCTGAGCATG | CTGGCAGCTG | TCTTGGAAGA |
| UNC5C8 | CAGAACTTCC | CAGATGGAAA | CCTGAGCATG | CTGGCAGCTG | TCTTGGAAGA |
| UNC5Cc | CAGAACTTCC | CAGATGGAAA | CCTGAGCATG | CTGGCAGCTG | TCTTGGAAGA |
| UNC5Cd | CAGAACTTCC | CAGATGGAAA | CCTGAGCATG | CTGGCAGCTG | TCTTGGAAGA |
| Consensus | CAGAACTTCC | CAGATGGAAA | CCTGAGCATG | CTGGCAGCTG | TCTTGGAAGA |

| | 1551 | | | | 1599 |
|---|---|---|---|---|---|
| UNC5C | AATGGGAAGA | CATGAAACGG | TGGTGTCCTT | AGCAGCAGAA | GGGCAGTAT |
| UNC5C8 | AATGGGAAGA | CATGAAACGG | TGGTGTCCTT | AGCAGCAGAA | GGGCAGTAT |
| UNC5Cc | AATGGGAAGA | CATGAAACGG | TGGTGTCCTT | AGCAGCAGAA | GGGCAGTAT |
| UNC5Cd | AATGGGAAGA | CATGAAACGG | TGGTGTCCTT | AGCAGCAGAA | GGGCAGTAT |
| Consensus | AATGGGAAGA | CATGAAACGG | TGGTGTCCTT | AGCAGCAGAA | GGGCAGTAT |

FIG. 2

```
Multalin version 5.3.3
Copyright I.N.R.A. France 1989, 1991, 1994, 1996
Published research using this software should cite
Multiple sequence alignment with hierarchical clustering
F. CORPET, 1988, Nucl. Acids Res., 16 22, 10881-10890
Symbol comparison table: blosum62
Gap weight: 12
Gap length weight: 2
Consensus levels: high=90% low=50%
Consensus symbols:
 ! is anyone of IV
 $ is anyone of LM
 % is anyone of FY
 # is anyone of NDQEBZ

 MSF:   2908    Check:    0
 Name: ratunc5h1      Len:   2908   Check: 8912   Weight:   0.87
 Name: ym97d12        Len:   2908   Check: 4745   Weight:   0.87
 Name: 1G             Len:   2908   Check: 1058   Weight:   1.05
 Name: 1Jrc           Len:   2908   Check:  508   Weight:   1.04
 Name: 2Brc           Len:   2908   Check: 6768   Weight:   1.04
 Name: 3D             Len:   2908   Check: 8193   Weight:   1.13
 Name: Consensus      Len:   2908   Check: 6031   Weight:   6.00

//

            1                                                  50
 ratunc5h1  ATGGCCGTCC GGCCCGGCCT GTGGCCAGTG CTCCTGGGCA TAGTCCTCGC
   ym97d12
        1G
      1Jrc
      2Brc
        3D
 Consensus

            51                                                100
 ratunc5h1  CGCCTGGCTT CGTGGTTCGG GTGCCCAGCA GAGTGCCACG GTGGCCAATC
   ym97d12
        1G
      1Jrc
      2Brc
        3D
 Consensus

            101                                               150
 ratunc5h1  CAGTGCCCGG TGCCAACCCC GACCTGCTGC CCCACTTCCT GGTAGAGCCT
   ym97d12
        1G
      1Jrc
      2Brc
        3D
 Consensus

            151                                               200
 ratunc5h1  GAGGACGTGT ACATTGTCAA GAACAAGCCG GTGTTGTTGG TGTGCAAGGC
   ym97d12
        1G
      1Jrc
      2Brc
        3D
 Consensus
```

FIG. 2 (CONTINUED 1).

```
            201                                                    250
ratunc5h1   TGTGCCTGCC ACCCAGATCT TCTTCAAGTG CAATGGGGAA TGGGTCCGCC
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            251                                                    300
ratunc5h1   AGGTCGATCA CGTAATTGAA CGCAGCACCG ACAGCAGCAG CGGATTGCCA
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            301                                                    350
ratunc5h1   ACCATGGAGG TCCGTATCAA CGTATCGAGG CAGCAGGTAG AGAAAGTGTT
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            351                                                    400
ratunc5h1   TGGGCTGGAG GAATACTGGT GCCAGTGTGT GGCATGGAGC TCCTCGGGTA
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            401                                                    450
ratunc5h1   CCACCAAAAG TCAGAAGGCC TACATCCGGA TTGCCTATTT GCGCAAGAAC
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            451                                                    500
ratunc5h1   TTTGAGCAGG AGCCACTGGC CAAGGAAGTG TCACTGGAGC AAGGCATTGT
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            501                                                    550
ratunc5h1   ACTACCTTGT CGCCCCCCAG AAGGAATCCC CCCAGCTGAG GTGGAGTGGC
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus
```

FIG. 2 (CONTINUED 2).

| | 551 | | | | 600 |
|---|---|---|---|---|---|
| ratunc5h1 | TTCGAAATGA | GGACCTCGTG | GACCCCTCCC | TCGATCCCAA | TGTGTACATC |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

| | 601 | | | | 650 |
|---|---|---|---|---|---|
| ratunc5h1 | ACGCGGGAGC | ACAGCCTAGT | CGTGCGTCAG | GCCCGCCTGG | CCGACACGGC |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

| | 651 | | | | 700 |
|---|---|---|---|---|---|
| ratunc5h1 | CAACTACACC | TGTGTGGCCA | AGAACATCGT | AGCCCGTCGC | CGAAGCACCT |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

| | 701 | | | | 750 |
|---|---|---|---|---|---|
| ratunc5h1 | CTGCAGCGGT | CATTGTTTAT | GTGAACGGTG | GGTGGTCGAC | GTGGACTGAG |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

| | 751 | | | | 800 |
|---|---|---|---|---|---|
| ratunc5h1 | TGGTCCGTCT | GCAGCGCCAG | CTGTGGGCGT | GGCTGGCAGA | AACGGAGCCG |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

| | 801 | | | | 850 |
|---|---|---|---|---|---|
| ratunc5h1 | GAGCTGCACC | AACCCGGCAC | CTCTCAACGG | GGGCGCCTTC | TGTGAGGGGC |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

| | 851 | | | | 900 |
|---|---|---|---|---|---|
| ratunc5h1 | AGAATGTCCA | GAAAACAGCC | TGCGCCACTC | TGTGCCCAGT | GGATGGGAGC |
| ym97d12 | | | | | |
| 1G | | | | | |
| 1Jrc | | | | | |
| 2Brc | | | | | |
| 3D | | | | | |
| Consensus | | | | | |

FIG. 2(CONTINUED 3).

```
            901                                                950
ratunc5h1   TGGAGTTCGT GGAGTAAGTG GTCAGCCTGT GGGCTTGACT GCACCCACTG
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            951                                               1000
ratunc5h1   GCGGAGCCGC GAGTGCTCTG ACCCAGCACC CCGCAATGGA GGTGAGGAGT
  ym97d12
       1G
     1Jrc
     2Brc
       3D
Consensus

            1001                                              1050
ratunc5h1   GTCGGGGTGC TGACCTGGAC ACCCGCAACT GTACCAGTGA CCTCTGCCTG
  ym97d12
       1G
     1Jrc
     2Brc
       3D                                       CAGTGA CCTCTGTGTA
Consensus

            1051                                              1100
ratunc5h1   CACACCGCTT CTTGCCCCGA GGACGTGGCT CTCTACATCG GCCTTGTCGC
  ym97d12
       1G
     1Jrc
     2Brc
       3D   CACACTGCTT CTGGCCCTGA GGACGTGGCC CTCTATGTGG GCCTNATCGC
Consensus

                        Predicted transmembrane region
            1101                                              1150
ratunc5h1   TGTGGCTGTG TGCCTCTTCT TGCTGTTGCT GGCCCTTGGA CTCATTTACT
  ym97d12
       1G
     1Jrc
     2Brc
       3D   CGTGGCCGNN TGCCTGGTCC TGCTGCTGCT TGTCCTCATC CTCGTTTATT
Consensus                  t  t     c    g      cc     c c c tt a

            1151                                              1200
ratunc5h1   GTCGCAAGAA GGAAGGGCTG GACTCCGATG TGGCCGACTC GTCCATCCTC
  ym97d12
       1G
     1Jrc
     2Brc
       3D   GCCGGAAGAA GGAGGGGCTG GACTCANATG TGGCTGACTC GTCCATTCTC
Consensus   gcc     aa gg        g ga        g t  c ga    c  t   t  tc

            1201                                              1250
ratunc5h1   ACCTCGGGCT TCCAGCCTGT CAGCATCAAG CCCAGCAAAG CAGACAACCC
  ym97d12
       1G
     1Jrc
     2Brc
       3D   ACCTCAGGCT TCCAGCCCGT CAGCATCAAG CCCAGCAAAG CAGACAACCC
Consensus    cc  a   t t   g cc t  agc   a        ca  g c      cc
```

FIG. 2 (CONTINUED 4).

```
            1251                                                  1300
ratunc5h1   CCACCTGCTC ACCATCCAGC CAGACCTCAG CACCACCACT ACCACCTACC
  ym97d12
       1G
     1Jrc   CCTTGGGTTC -CCNTCAAGT GGTNNCANGG GGGTGGCCCT TGAA--TTCA
     2Brc   ACTTGGGTTC -CCNTCAAGT TGT--CAATG GGGNGCCCCT --GA--ATCA
       3D   CCATCTGCTC ACCATCCAGC CGGACCTCAG CACCACCACC ACCACCTACC
Consensus   cc t  g tc  cc tc ag   g   c   g      cc c     a  t c

            1301                                                  1350
ratunc5h1   AGGGCAGTCT ATGTTCGAGG CAGGATGGAC CCAGCCCCAA GTTCCAGCTC
  ym97d12
       1G
     1Jrc
     2Brc
       3D   AGGGCAGTCT NTGTCCCCGG CAGGATGGGC CCAGCCCCAA GTTCCAGCTC
Consensus   ag   a t    tgt    gg   gg tgg   c agc c       ccag

            1351                                                  1400
ratunc5h1   TCTAATGGTC ACCTGCTCAG CCCACTGGGG AGTGGCCGCC ATACGTTGCA
  ym97d12                                           GCC ACAC--TGCA
       1G                    TCAG CCCCCTGGGT GGCGGCCGCC ACACACTGCA
     1Jrc
     2Brc
       3D   ACCAATGGGC ACCTGCTCAG CCCCCTGGGT GGCGGCCGCC ACACACTGCA
Consensus      aa g  c  cct  tcag ccc cctggg  g ggccgCC acac  tGCA

            1401                                                  1450
ratunc5h1   CCACAGCTCA CCCACCTCTG AGGCTGAGGA CTTCGTCTCC CGCCTCTCCA
  ym97d12   CCACAGCTCT CCCACCTCTG AGGCCGAGGA GTTCGTCTCC CGCCTCTCCA
       1G   CCACAGCTCT CCCACCTCTG AGGCCGAGGA GTTCGTCTCC CGCCTCTCCA
     1Jrc
     2Brc
       3D   CCACAGCTCT CCAACCTNTG AGGCCNAGGA GTTCGNNTCC CGCCTTTCCA
Consensus   cCacagCtct cCcacctctG aggcc AGGa gttCg  tcc cGccT Tcca

            1451                                                  1500
ratunc5h1   CCCAAAACTA CTT-TCGTTC CCTGCCCCGC GGCACCAGCA ACATGGCCTA
  ym97d12   CCCAGAACTA CTT-CCGCTC CCTGCCCCGA GGCACCAGCA ACATGACCTA
       1G   CCCAGAACTA CTT-CCGCTC CCTGCCCCGA GGCACCAGCA ACATGACCTA
     1Jrc
     2Brc
       3D   CCCAGAACTA CTTNCGGTTC CTTGCCCCCA GGCNCCAGCA ACATGACCTT
Consensus   cccagaacTa ctT cgGttC ctTgccCcga GGc ccagca acAtGaCCT

            1501                                                  1550
ratunc5h1   C--GGGACCT TCA-ACTTCC TCGGGGG-CC GGCTGATGAT --CCCTAATA
  ym97d12   T--GGGACCT TCA-ACTTCC TCGGGGG-CC GGCTGATGAT --CCCTAATA
       1G   T--GGGACCT TCNNACTTCC TCGGGGG-CC GGCTGATGAT --CCCTAATA
     1Jrc
     2Brc
       3D   ATGGGGACCT TTAAATTTCT TCGGGGGNCC GGNTTATGAA NCCCTAATTC
Consensus      gGGaCCT t   acTTCc TcggggG CC Gg t atga    cc  atTc

            1551                                                  1600
ratunc5h1   CGGGGA--TC AGCCTCCT-C ATACCCCCGG ATGCCATCCC CC-GAGGAAA
  ym97d12   CAGGAA--TC AGCCTCCT-C ATCCCCCCAG ATGCCATACC CC-GAGGGAA
       1G   CAGGAA--TC AGCCTCCT-C ATNCCCCCAG ATGCCATACC CC-GAGGGAA
     1Jrc   CAGGAA--TC AGCCTCCT-C ATCCCCCCAG ATGCCATACC CC-GAGGGAA
     2Brc   CAGGAA--TC AGCCTCCT-C ATCCCCCCAG ATGCCATACC CC-GAGGGAA
       3D   CAGGGAATTA AACCTTCTTA ATCCCCCCAA ATGCCANACC CCCGANGGAA
Consensus   CaGGaA  Tc AgCCTcCT c ATcCCCCCag ATGCCAtaCC CC GAgGgAA
```

FIG. 2 (CONTINUED 5).

```
           1601                                                1650
ratunc5h1  GATCT-ACGA GATCTACCTC ACACTGCACA AGCCAGAAGA CGTGAGGTTG
ym97d12    GATCT-ATGA GATCTACCTC ACGCTGCACA AGCCGGAAGA CGTGAGGTTG
1G         GATCT-ATGA GATCTGCCTC ACGCTGCACA AGCCGGAAGA CGTGAGGTTG
1Jrc       GATCT-ATGA GATCTACCTC ACGCTGCACA AGCCGGAAGA CGTGAGGTTG
2Brc       GATCT-ATGA GATCTACCTC ACGCTGCGCA AGCCGGAAGA CGTGAGGTTG
3D         NATCTNTTGN NAACTACCTT A--------A ANCTTGANNA AGCCCGGAAA
Consensus  gATCT atGa gAtCTaCCTc AcgctgcacA AgCcgGAagA cGtgaGGttg

           1651                                                1700
ratunc5h1  CCCCTAGCTG GCTGTCAGAC CCTGCTGAGT CCAGTCGTTA GCTGTGGGCC
ym97d12    CCCCTAGCTG GCTGTCAGAC CCTGCTGAGT CCCATCGTTA GCTGTGGACC
1G         CCCCTAGCTG GCTGTCAGAC CCTGCTGAGT CCCATCGTTA GCTGTGGACC
1Jrc       CCCCTAGCTG GCTGTCAGAC CCTGCTGAGT CCCATCGTTA GCTGTGGACC
2Brc       CCCCTAGCTG GCTGTCAGAC CCTGCTGAGT CCCATCGTTA GCTGTGGACC
3D         AACC
Consensus  cccctagctg gctgtcagac cctgctgagt cccatcgtta gctgtggacc

           1701                                                1750
ratunc5h1  CCCA-GGAGT CCTGCTCACC CGGCCAGTCA T-CCTTG-CA ATGGACCACT
ym97d12    CCCT-GGCGT CCTGCTCACC CGGCCAGTCA T-CCTGG-CT ATGGACCACT
1G         CCCT-GGCGT CCTGCTCACC CGGCCAGTCA T-CCTGG-CT ATGGACCACT
1Jrc       CCCT-GGCGT CCTGCTCACC CGGCCAGTCA T-CCTGG-CT ATGGACCACT
2Brc       CCCT-GGCGT CCTGCTCACC CGGCCAGTCA T-CCTGG-CT ATGGACCACT
3D
Consensus  ccct ggcgt cctgctcacc cggccagtca t cctgg ct atggaccact

           1751                                                1800
ratunc5h1  GT--GGAGAG CCCA-GCCCT -GACAGC--T GGAGTC-TGC GCCT---CAA
ym97d12    GT--GGGGAG CCCA-GCCCT -GACAGC--T GGAGCC-TGC GCCT---CAA
1G         GT--GGGGAG CCCA-GCCCT -GACAGC--T GGAGCC-TGC GCCT---CAA
1Jrc       GT--GGGGAG CCCA-GCCCT -GACAGC--T GGAGCC-TGC GCCT---CAA
2Brc       GT--GGGGAG CCCA-GCCCT -GACAGC--T GGAGCC-TGC GCCT---CAA
3D
Consensus  gt  ggggag ccca gccct  gacagc  t ggagcc tgc gcct   caa

           1801                                                1850
ratunc5h1  AAAGCAG-TC CTGC-GAGGG CAGTTGGG-- -AGGATGTGC -TGCACCT-T
ym97d12    AAAGCAG-TC GTGC-GAGGG CAGCTGGG-- -AGGATGTGC -TGCACCT-G
1G         AAAGCAG-TC GTGC-GAGGG CAGCTGGG-- -AGGGTGTGC -TGCACCT-G
1Jrc       AAAGCAG-TC GTGC-GAGGG CAGCTGGG-- -AGGATGTGC -TGCACCT-G
2Brc       AAAGCAG-TC GTGC-GAGGG CAGCTGGG-- -AGGATGTGC -TGCACCT-G
3D
Consensus  aaagcag tc  tgc gaggg cagctggg    aggatgtgc  tgcacct g

           1851                                                1900
ratunc5h1  GGTGAGGAGT CACCTTCCCA CCTCTACTAC TGCCAGCTGG AGGCCGGGGC
ym97d12    GGCGAGGAGG CGCCCTCCCA CCTCTACTAC TGCCAGCTGG AGGCCAGTGC
1G         GGCGAGGAGG CGCCCTCCCA CCTCTACTAA NTAAANCCCN AA-TTNTTGC
1Jrc       GGCGAGGAGG CGCCCTCCCA CCTCTACTAG
2Brc       GGCGAGGAGG CGCCCTCCCA CCTCTACTAA G
3D
Consensus  ggcgaggag  cgccctccca cctctacta

           1901                                                1950
ratunc5h1  CTGCTATGTC TTCACGGAGC AGCTGGGCCG CTTTGCCCTG GTAGGAGAGG
ym97d12    CTGCTACGTC TTCACCGAGC AGCTGGGCCG CTTTGCCCTG GTGGGAGAGG
1G         AAAAATCCNT TTAAAATTGT NG--GNCCCN TTNAAACCTN ----------
1Jrc
2Brc
3D
Consensus
```

FIG. 2 (CONTINUED 6).

```
             1951                                                    2000
ratunc5h1    CCCTCAGCGT GGCTGCCACC AAGCGCCTCA GGCTCCTTCT GTTTGCTCCC
ym97d12      CCCTCAGCGT GGCTGCCGCC AAGCGCCTCA AGCTGCTTCT GTTTGCGCCG
1G           CCCTTAAAAA GGGGCCCAAT TTCCNCCTNT NNGGNANCCN --TTNAAAAN
1Jrc
2Brc
3D
Consensus

             2001                                                    2050
ratunc5h1    GTGGCCTGTA CGTCCCTTGA GTACAACATC CGAGTGTACT GCCTACACGA
ym97d12      GTGGCCTGCA CCTCCCTCGA GTACAACATC CGGGTCTACT GCCTGCATGA
1G           NTAACTGGCC CCTNTTTTNA AAACNNNCGA NCNGGGNAAA NCC
1Jrc
2Brc
3D
Consensus

             2051                                                    2100
ratunc5h1    CACCCACGAC GCTCTCAAGG AGGTGGTGCA GCTGGAGAAG CAGCTAGGTG
ym97d12      CACCCACGAT GCACTCAAGG AGGTGGTGCA GCTGGAGAAG CAGCTGGGGG
1G
1Jrc
2Brc
3D
Consensus

             2101                                                    2150
ratunc5h1    GACAGCTGAT CCAGGAGCCT CGCGTCCTGC ACTTCAAAGA CAGTTACCAC
ym97d12      GACAGCTGAT CCAGGAGCCA CGGGTCCTGC ACTTCAAGGA CAGTTACCAC
1G
1Jrc
2Brc
3D
Consensus

             2151                                                    2200
ratunc5h1    AACCTACGTC TCTCCATCCA CGACGTGCCC AGCTCCCTGT GGAAGAGCAA
ym97d12      AACCTGCGCC TATCCATCCA CGATGTGCCC AGCTCCCTGT GGAAGAGTAA
1G
1Jrc
2Brc
3D
Consensus

             2201                                                    2250
ratunc5h1    GCTACTTGTC AGCTACCAGG AGATCCCTTT TTACCACATC TGGAACGGCA
ym97d12      GCTCCTTGTC AGCTACCAGG AGATCCCCTT TTATCACATC TGGAATGGCA
1G
1Jrc
2Brc
3D
Consensus

             2251                                                    2300
ratunc5h1    CCCAGCAGTA TCTGCACTGC ACCTTCACCC TGGAGCGCAT CAACGCCAGC
ym97d12      CGCAGCGGTA CTTGCACTGC ACCTTCACCC TGGAGCGTGT CAGCCCCAGC
1G
1Jrc
2Brc
3D
Consensus
```

FIG. 2 (CONTINUED 7).

```
          2301                                                2350
ratunc5h1 ACCAGCGACC TGGCCTGCAA GGTGTGGGTG TGGCAGGTGG AGGGAGATGG
  ym97d12 ACTAGTGACC TGGCCTGCAA GCTGTGGGTG TGGCAGGTGG AGGGCGACGG
       1G
     1Jrc
     2Brc
       3D
Consensus

          2351                                                2400
ratunc5h1 GCAGAGCTTC AACATCAACT TCAACATCAC TAAGGACACA AGGTTTGCTG
  ym97d12 GCAGAGCTTC AGCATCAACT TCAACATCAC CAAGGACACA AGGTTTGCTG
       1G
     1Jrc
     2Brc
       3D
Consensus

          2401                                                2450
ratunc5h1 AATTGTTGGC TCTGGAGAGT GAAGGGGGGG TCCCAGCCCT GGTGGGCCCC
  ym97d12 AGCTGCTGGC TCTGGAGAGT GAAGCGGGGG TCCAAGCCCT GGTGGGCCCC
       1G
     1Jrc
     2Brc
       3D
Consensus

          2451                                                2500
ratunc5h1 AGTGCCTTCA AGATCCCCTT CCTCATTCGG CAAAAGATCA TCGCCAGTCT
  ym97d12 AGTGCCTTCA AGATCCCCTT CCTCATTCGG CAGAAGATAA TTTCCAGCCT
       1G
     1Jrc
     2Brc
       3D
Consensus

          2501                                                2550
ratunc5h1 GGACCCACCC TGCAGCCGGG GCGCCGACTG GAGAACTCTA GCCCAGAAAC
  ym97d12 GGACCCACCC TGTAGGCGGG GTGCCGACTG GCGGACTCTG GCCCAGAAAC
       1G
     1Jrc
     2Brc
       3D
Consensus

          2551                                                2600
ratunc5h1 TTCACCTGGA CAGCCATCTT AGCTTCTTTG CCTCCAAGCC CAGCCCTACA
  ym97d12 TCCACCTGGA CAGCCATCTC AGCTTCTTTG CCTCCAAGCC CAGCCCCACA
       1G
     1Jrc
     2Brc
       3D
Consensus

          2601                                                2650
ratunc5h1 GCCATGATCC TCAACCTATG GGAGGCACGG CACTTCCCCA ACGGCAACCT
  ym97d12 GCCATGATCC TCAACCTGTG GGAGGCGCGG CACTTCCCCA ACGGCAACCT
       1G
     1Jrc
     2Brc
       3D
Consensus
```

FIG. 2 (CONTINUED 8).

```
             2651                                                    2700
ratunc5h1    CGGCCAGCTG GCAGCAGCTG TGGCCGGACT GGGCCAACCA GATGCTGGCC
  ym97d12    CAGCCAGCTG GCTGCAGCAG TGGCTGGACT GGGCCAGCCA GACGCTGGCC
       1G
     1Jrc
     2Brc
       3D
Consensus

             2701                                                    2750
ratunc5h1    TCTTCACGGT GTCGGAGGCC GAGTGTTGA
  ym97d12    TCTTCACAGT GTCGGAGGCT GAGTGCTGAG GCCGGCCAGG CCCGACACCT
       1G
     1Jrc
     2Brc
       3D
Consensus

             2751                                                    2800
ratunc5h1
  ym97d12    ACACTCTCAC CAGCTTTGGC ACCCACCAAG GACAGGCAGA AGCCGGACAG
       1G
     1Jrc
     2Brc
       3D
Consensus

             2801                                                    2850
ratunc5h1
  ym97d12    GGGCCCTTCC CCACACCGGG GAGAGCTGCT CGGACAGGCC CCCTCCCGGC
       1G
     1Jrc
     2Brc
       3D
Consensus

             2851                                                    2900
ratunc5h1
  ym97d12    CGAAGCTGTC CCTTAATGCT GGTCCTTCAG ACCCTGCCCC CTCGTGCCGA
       1G
     1Jrc
     2Brc
       3D
Consensus

             2901
ratunc5h1
  ym97d12    ATTCTGGC
       1G
     1Jrc
     2Brc
       3D
Consensus
```

FIG. 3.

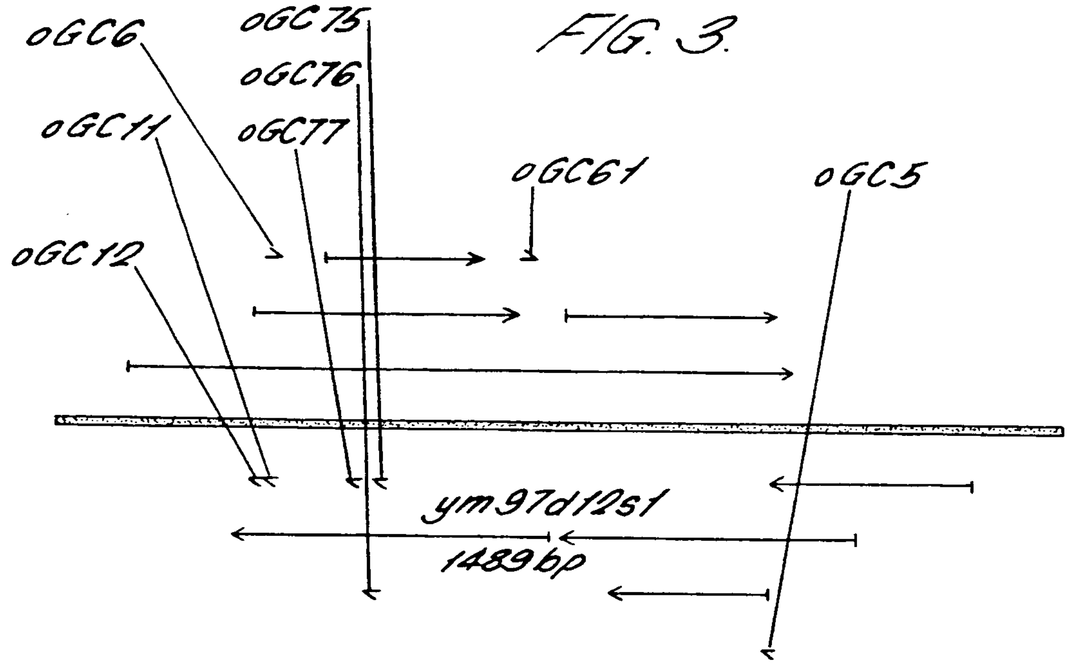

oGC6
oGC75
oGC76
oGC77
oGC11
oGC61
oGC5
oGC12
ym97d12s1
1489bp

FIG. 4.

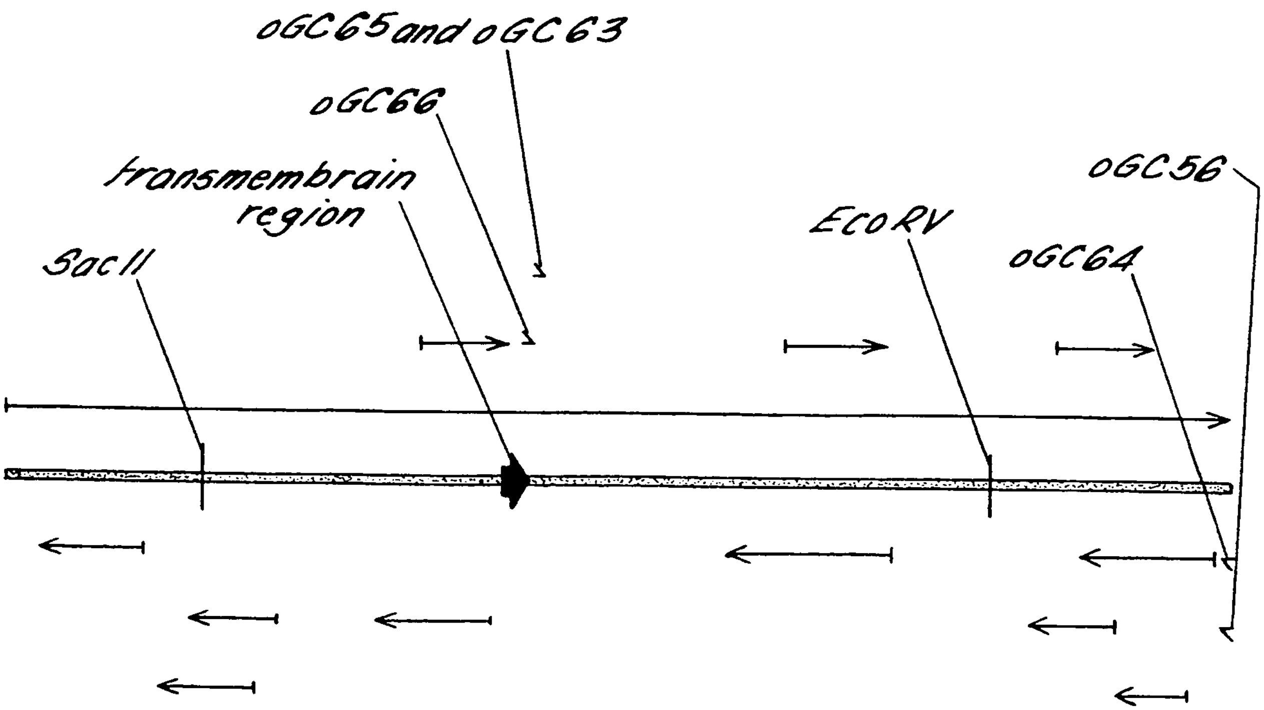

oGC65 and oGC63
oGC66
transmembrain region
SacII
EcoRV
oGC56
oGC64

FIG. 5.

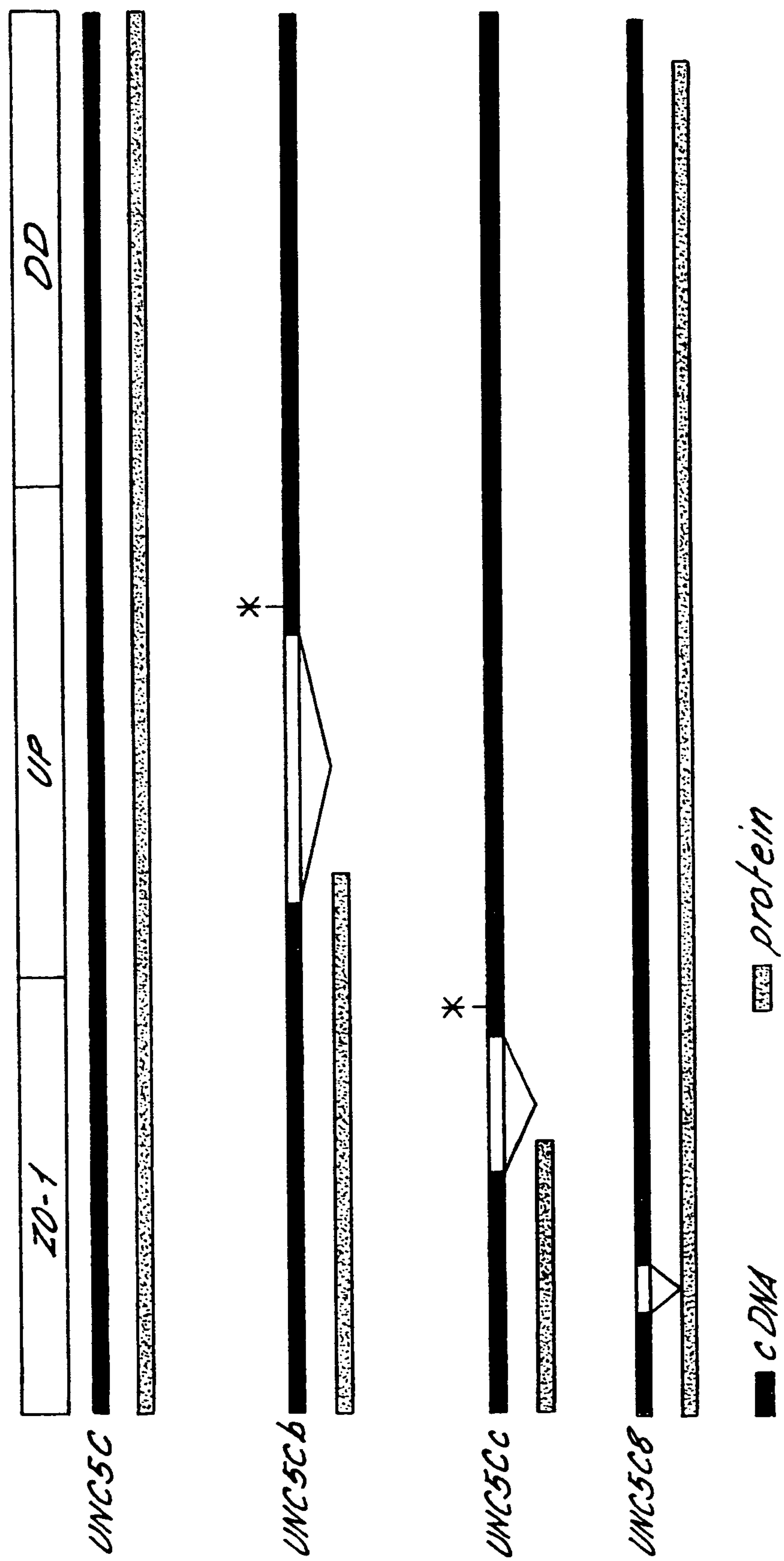


FIG. 6.

gi|205715 (M96376) neurexin II-alpha-b [Rattus norvegicus] 31 7.4

```
 gi|205715 (M96376) neurexin II-alpha-b [Rattus norvegicus]
           Length = 1728

 Score = 31.3 bits (69), Expect = 7.4
 Identities = 16/38 (42%), Positives = 20/38 (52%)

Query: 337  KACSVCXAGRRALMGKLLEEQGXGVGGRGKANADIYYR 224
            KAC VC       + GK LEE+G G  G G+    IY +
Sbjct: 1690 KACCVCRCRATCIAGKPLEERGGG-RGEGERQMQIYIK 1726
```

FIG. 7.

```
gi|1644455 (U72520) mena protein [Mus musculus]
          Length = 541

 Score = 34.0 bits (76), Expect = 0.77
 Identities = 14/23 (60%), Positives = 15/23 (64%)
 Frame = +1

Query: 31  PPPPCTCPAGRHRVSALPPPAGP 99
           PPPP   P+G    SALPPP GP
Sbjct: 284 PPPPPPLPSGPAYASALPPPPGP 306
```

FIG. 8.

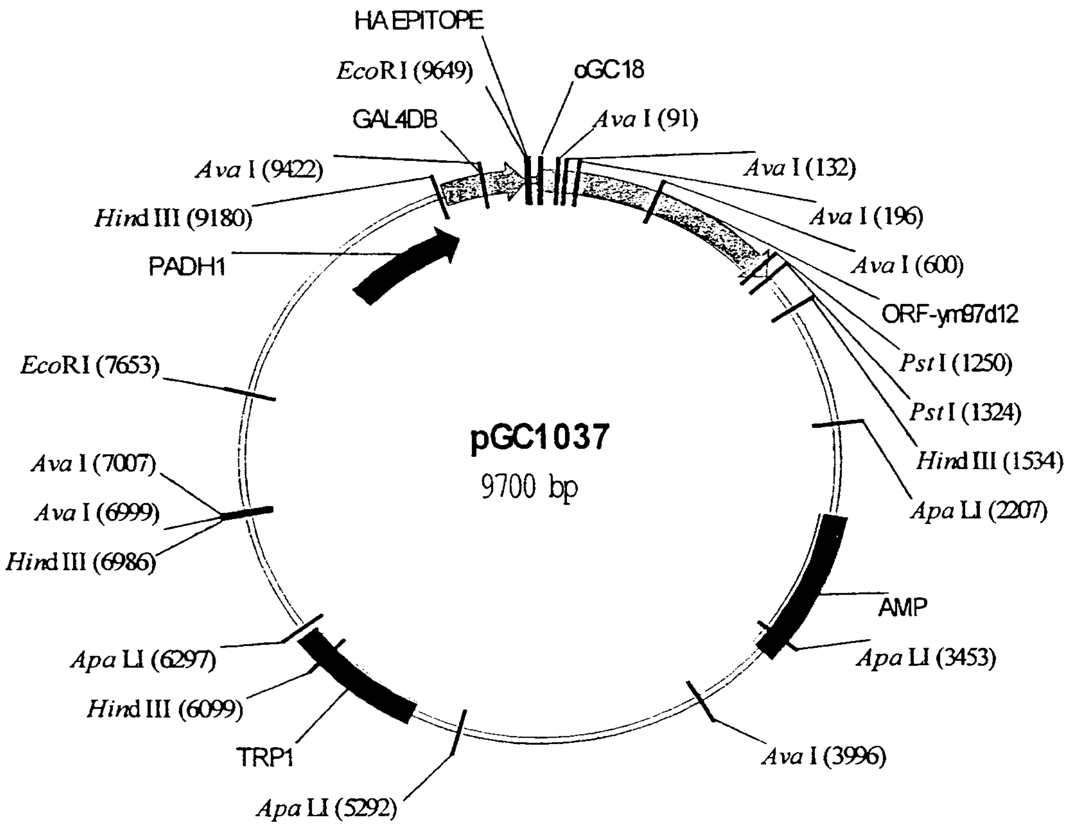
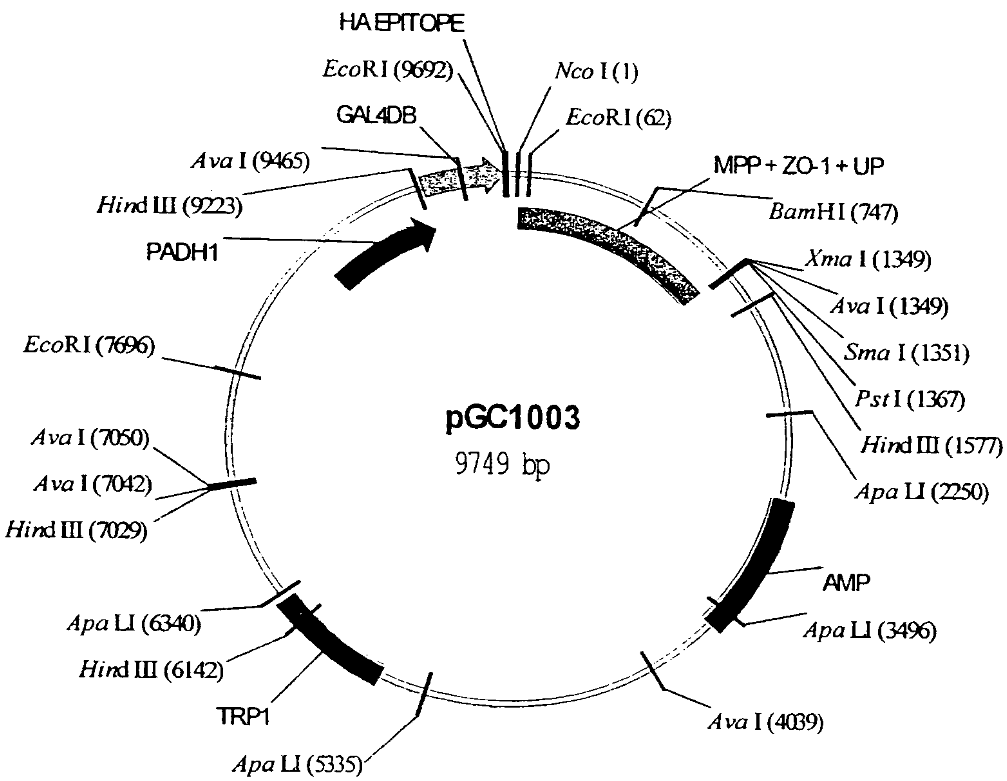

HA EPITOPE
EcoR I (9649)
GAL4DB
Ava I (9422)
Hind III (9180)
PADH1
EcoR I (7653)
Ava I (7007)
Ava I (6999)
Hind III (6986)
Apa LI (6297)
Hind III (6099)
TRP1
Apa LI (5292)
cGC18
Ava I (91)
Ava I (132)
Ava I (196)
Ava I (600)
ORF-ym97d12
Pst I (1250)
Pst I (1324)
Hind III (1534)
Apa LI (2207)
AMP
Apa LI (3453)
Ava I (3996)
pGC1037
9700 bp

FIG. 9.


HA EPITOPE
EcoRI (9692)
Nco I (1)
EcoRI (62)
GAL4DB
Ava I (9465)
MPP + ZO-1 + UP
Hind III (9223)
BamHI (747)
PADH1
Xma I (1349)
Ava I (1349)
Sma I (1351)
EcoRI (7696)
Pst I (1367)
pGC1003
9749 bp
Hind III (1577)
Ava I (7050)
Apa LI (2250)
Ava I (7042)
Hind III (7029)
AMP
Apa LI (6340)
Apa LI (3496)
Hind III (6142)
TRP1
Ava I (4039)
Apa LI (5335)

SEQUENCE LISTING

<110> DEVGEN NV

<120> UNC-5 constructs and screening methods

<130> SCB/52877/002

<140>
<141>

<160> 96

<170> PatentIn Ver. 2.0

<210> 1
<211> 1393
<212> DNA
<213> Homo sapiens

<400> 1
ttgtttgtgt atcggaagaa tcatcgtgac tttgagtcag atattattga ctcttcggca 60
ctcaatgggg gctttcagcc tgtgaacatc aaggcagcaa gacaagatct gctggctgta 120
cccccagacc tcacgtcagc tgcagccatg tacagaggac ctgtctatgc cctgcatgac 180
gtctcagaca aaatcccaat gaccaactct ccaattctgg atccactgcc caacctgaaa 240
atcaaagtgt acaacacctc aagtgctgtc accccccaag atgacctctc tgagtttacg 300
tccaagctgt cccctcagat gacccagtcg ttgttggaga atgaagccct cagcctgaag 360
aaccagagtc tagcaaggca gactgatcca tcctgtaccg catttggcag cttcaactcg 420
ctgggaggtc accttattgt tcccaattca ggagtcagct tgctgattcc cgctggggcc 480
attccccaag ggagagtcta cgaaatgtat gtgactgtac acaggaaaga aactatgagg 540
ccacccatgg atgactctca gacacttttg acccctgtgg tgagctgtgg gcccccagga 600
gctctgctca cccgccccgt cgtcctcact atgcatcact gcgcagaccc caataccgag 660
gactggaaaa tactgctcct cgctggagta cagcatccga gtctactgtc tggatgacac 720
ccaggatgcc ctgaaggaaa ttttacatct tgagagacag acgggaggac agctcctaga 780
agaacctaag gctcttcatt ttaaaggcag cacccacaac ctgcgcctgt caattcacga 840
tatcgcccat tccctctgga agagcaaatt gctggctaaa tatcaggaaa ttccatttta 900
ccatgtttgg agtggatctc aaagaaacct gcactgcacc ttcactctgg aaagatttag 960
cctgaacaca gtggagctgg tttgcaaact ctgtgtgcgg caggtggaag gagaagggca 1020
gatcttccag ctcaactgca ccgtgtcaga ggaacctact ggcatcgatt tgccgctgct 1080
ggatcctgcg aacaccatca ccacggtcac ggggcccagt gctttcagca tccctctccc 1140
tatccggcag aagctctgta gcagcctgga tgccccccag acgagaggcc atgactggag 1200
gatgctggcc cataagctga acctggacag gtacttgaat tactttgcca ccaaatccag 1260
cccaactggc gtaatcctgg atctttggga agcacagaac ttcccagatg gaaacctgag 1320
catgctggca gctgtcttgg aagaaatggg aagacatgaa acggtggtgt ccttagcagc 1380
agaagggcag tat 1393

<210> 2
<211> 238
<212> PRT
<213> Homo sapiens

<400> 2
Leu Phe Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile
1 5 10 15

Asp Ser Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala
20 25 30

```
Ala Arg Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala
        35                  40                  45

Ala Met Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys
    50                  55                  60

Ile Pro Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys
65                  70                  75                  80

Ile Lys Val Tyr Asn Thr Ser Ser Ala Val Thr Pro Gln Asp Asp Leu
                85                  90                  95

Ser Glu Phe Thr Ser Lys Leu Ser Pro Gln Met Thr Gln Ser Leu Leu
            100                 105                 110

Glu Asn Glu Ala Leu Ser Leu Lys Asn Gln Ser Leu Ala Arg Gln Thr
        115                 120                 125

Asp Pro Ser Cys Thr Ala Phe Gly Ser Phe Asn Ser Leu Gly Gly His
    130                 135                 140

Leu Ile Val Pro Asn Ser Gly Val Ser Leu Leu Ile Pro Ala Gly Ala
145                 150                 155                 160

Ile Pro Gln Gly Arg Val Tyr Glu Met Tyr Val Thr Val His Arg Lys
                165                 170                 175

Glu Thr Met Arg Pro Pro Met Asp Asp Ser Gln Thr Leu Leu Thr Pro
            180                 185                 190

Val Val Ser Cys Gly Pro Pro Gly Ala Leu Leu Thr Arg Pro Val Val
        195                 200                 205

Leu Thr Met His His Cys Ala Asp Pro Asn Thr Glu Asp Trp Lys Ile
    210                 215                 220

Leu Leu Leu Ala Gly Val Gln His Pro Ser Leu Leu Ser Gly
225                 230                 235


<210> 3
<211> 1438
<212> DNA
<213> Homo sapiens

<400> 3
ttgtttgtgt atcggaagaa tcatcgtgac tttgagtcag atattattga ctcttcggca 60
ctcaatgggg gctttcagcc tgtgaacatc aaggcagcaa gacaagatct gctggctgta 120
cccccagacc tcacgtcagc tgcagccatg tacagaggac ctgtctatgc cctgcatgac 180
gtctcagaca aaatcccaat gaccaactct ccaattctgg atccactgcc caacctgaaa 240
atcaaagtgt acaacacctc aggtgctgtc acctattgtt cccaattcag gagtcagctt 300
gctgattccc gctggggcca ttccccaagg gagagtctac gaaatgtatg tgactgtaca 360
caggaaagaa actatgaggc cacccatgga tgactctcag acacttttga cccctgtggt 420
gagctgtggg cccccaggag ctctgctcac ccgccccgtc gtcctcacta tgcatcactg 480
cgcagacccc aataccgagg actggaaaat actgctcaag aaccaggcag cacagggaca 540
gtgggaggat gtggtggtgg tcggggagga aaacttcacc accccctgct acattcagct 600
ggatgcagag gcctgccaca tcctcacaga gaacctcagc acctacgccc tggtaggaca 660
```

ttccaccacc aaagcggctg caaagcgcct caagctggcc atctttgggc ccctgtgctg 720
ctcctcgctg gagtacagca tccgagtcta ctgtctggat gacacccagg gtgccctgaa 780
ggaaatttta catcttgaga gacagacggg aggacagctc ctagaagaac ctaaggctct 840
tcattttaaa ggcagcaccc acaacctgcg cctgtcaatt cacgatatcg cccgttccct 900
ctggaagagc aaattgctgg ctaaatatca ggaaattcca ttttaccatg tttggagtgg 960
atctcaaaga aacctgcact gcaccttcac tctggaaaga tttagcctga acacagtgga 1020
gctggtttgc aaactctgtg tgcggcaggt ggaaggagaa gggcagatct tccagctcaa 1080
ctgcaccgtg tcagaggaac ctactggcat cgatttgccg ctgctggatc ctgcgaacac 1140
catcaccacg gtcacggggc ccagtgcttt cagcatccct ctccctatcc ggcagaagct 1200
ctgtagcagc ctggatgccc cccagacgag aggccatgac tggaggatgc tggcccataa 1260
gctgaacctg gacaggtact tgaattactt tgccaccaaa tccagcccaa ctggcgtaat 1320
cctggatctt tgggaagcac agaacttccc agatggaaac ctgagcatgc tggcagctgt 1380
cttggaagaa atgggaagac atgaaacggt ggtgtcctta gcagcagaag ggcagtat 1438

<210> 4
<211> 130
<212> PRT
<213> Homo sapiens

<400> 4
Leu Phe Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile
1 5 10 15

Asp Ser Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala
20 25 30

Ala Arg Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala
35 40 45

Ala Met Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys
50 55 60

Ile Pro Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys
65 70 75 80

Ile Lys Val Tyr Asn Thr Ser Gly Ala Val Thr Tyr Cys Ser Gln Phe
85 90 95

Arg Ser Gln Leu Ala Asp Ser Arg Trp Gly His Ser Pro Arg Glu Ser
100 105 110

Leu Arg Asn Val Cys Asp Cys Thr Gln Glu Arg Asn Tyr Glu Ala Thr
115 120 125

His Gly
130

<210> 5
<211> 1575
<212> DNA
<213> Homo sapiens

<400> 5
ttgtttgtgt atcggaagaa tcatcgtgac tttgagtcag atattattga ctcttcggca 60
ctcaatgggg gctttcagcc tgtgaacatc aaggcagcaa gacaagacct cacgtcagct 120
gcagccatgt acagaggacc tgtctatgcc ctgcatgacg tctcagacaa aatcccaatg 180
accaactctc caattctgga tccactgccc aacctgaaaa tcaaagtgta caacacctca 240

```
ggtgctgtca ccccccaaga tgacctctct gagtttacgt ccaagctgtc ccctcagatg 300
acccagtcgt tgttggagaa tgaagccctc agcctgaaga accagagtct agcaaggcag 360
actgatccat cctgtaccgc atttggcagc ttcaactcgc tgggaggtca ccttattgtt 420
cccaattcag gagtcagctt gctgattccc gctggggcca ttccccaagg gagagtctac 480
gaaatgtatg tgactgtaca caggaaagaa actatgaggc cacccatgga tgactctcag 540
[illegible] cccctgtggt gagctgtggg cccccaggag ctctgctcac ccgcccc gcc 600
gtcttcacta tgcatcactg cgcagacccc aataccgagg actggaaaat actgctcaag 660
[illegible] cacagggaca gtgggaggat gtggtggtgg ccggggagga aaacttcacc 720
[illegible] acattaagct ggatgcagag gcctgccaca tcctcacaga gaacctcagc 780
[illegible] tggtaggaca ttccaccacc aaagcggctg caaagcgcct caagctggcc 840
[illegible] ccctgtgctg ctcctcgctg gagtacagca tccgagtcta ctgtctggat 900
[illegible] atgccctgaa ggaaatttta catcttgaga gannnnnnnn nnnnnnnnnn 960
nnnnnnnnnn nnnaggtttt tcatttaaag cangcanccn ncaaatgngc ctgtcaattc 1020
[illegible] ccattccctc tgaaagagca aattgctggc taaatatcag gaaattccat 1080
[illegible] ttggagtgga tctcaaagaa ncctgcactg cacnttcact ctggaaagat 1140
[illegible] cacagtggag ctggtttgca aactctgtgt cggcaggtgg aaggagaagg 1200
[illegible] agctcaactg caccgtgtca gaggaaccta ctggcatcga tttgccgctg 1260
[illegible] cgaacaccat caccacggtc acggggccca gtgctttcag catccctctc 1320
[illegible] agaagctctg tagcagcctg gatgcccccc agacgagagg ccatgactgg 1380
[illegible] cccataagct gaacctggac gggtacttga attactttgc caccaaatcc 1440
[illegible] acgtaatcct ggatctttgg gaagcacaga acttcccaga tggaaacctg 1500
[illegible] cagctgtctt ggaagaaatg ggaagacatg aaacggtggt gtccttagca 1560
[illegible] agtat                                                   1575
```

```
<210> 6
<211> 526
<212> PRT
<213> Homo sapiens

<400> 6
Leu Phe Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile
1               5                   10                  15

Asp Ser Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala
            20                  25                  30

Ala Arg Gln Asp Leu Thr Ser Ala Ala Ala Met Tyr Arg Gly Pro Val
        35                  40                  45

Tyr Ala Leu His Asp Val Ser Asp Lys Ile Pro Met Thr Asn Ser Pro
    50                  55                  60

Ile Leu Asp Pro Leu Pro Asn Leu Lys Ile Lys Val Tyr Asn Thr Ser
65                  70                  75                  80

Gly Ala Val Ser Pro Gln Asp Asp Leu Ser Glu Phe Thr Ser Lys Leu
                85                  90                  95

Ser Pro Gln Met Thr Gln Ser Leu Leu Glu Asn Glu Ala Leu Ser Leu
            100                 105                 110

Lys Asn Gln Ser Leu Ala Arg Gln Thr Asp Pro Ser Cys Thr Ala Phe
        115                 120                 125

Gly Ser Phe Asn Ser Leu Gly Gly His Leu Ile Val Pro Asn Ser Gly
    130                 135                 140

Val Ser Leu Leu Ile Pro Ala Gly Ala Ile Pro Gln Gly Arg Val Tyr
```

145 150 155 160

Glu Met Tyr Val Thr Val His Arg Lys Glu Thr Met Arg Pro Pro Met
165 170 175

Asp Asp Ser Gln Thr Leu Leu Thr Pro Val Val Ser Cys Gly Pro Pro
180 185 190

Gly Ala Leu Leu Thr Arg Pro Val Val Leu Thr Met His His Cys Ala
195 200 205

Asp Pro Asn Thr Glu Asp Trp Lys Ile Leu Leu Lys Asn Gln Ala Ala
210 215 220

Gln Gly Gln Trp Glu Asp Val Val Val Val Gly Glu Glu Asn Phe Thr
225 230 235 240

Thr Pro Cys Tyr Ile Lys Leu Asp Ala Glu Ala Cys His Ile Leu Thr
245 250 255

Glu Asn Leu Ser Thr Tyr Ala Leu Val Gly His Ser Thr Thr Lys Ala
260 265 270

Ala Ala Lys Arg Leu Lys Leu Ala Ile Phe Gly Pro Leu Cys Cys Ser
275 280 285

Ser Leu Glu Tyr Ser Ile Arg Val Tyr Cys Leu Asp Asp Thr Gln Asp
290 295 300

Ala Leu Lys Glu Ile Leu His Leu Glu Arg Gln Thr Gly Gly Gln Leu
305 310 315 320

Leu Glu Glu Pro Lys Ala Leu His Phe Lys Gly Ser Thr His Asn Leu
325 330 335

Arg Leu Ser Ile His Asp Ile Ala His Ser Leu Trp Lys Ser Lys Leu
340 345 350

Leu Ala Lys Tyr Gln Glu Ile Pro Phe Tyr His Val Trp Ser Gly Ser
355 360 365

Gln Arg Asn Leu His Cys Thr Phe Thr Leu Glu Arg Phe Ser Leu Asn
370 375 380

Thr Val Glu Leu Val Cys Lys Leu Cys Val Arg Gln Val Glu Gly Glu
385 390 395 400

Gly Gln Ile Phe Gln Leu Asn Cys Thr Val Ser Glu Glu Pro Thr Gly
405 410 415

Ile Asp Leu Pro Leu Leu Asp Pro Ala Asn Thr Ile Thr Thr Val Thr
420 425 430

Gly Pro Ser Ala Phe Ser Ile Pro Leu Pro Ile Arg Gln Lys Leu Cys
435 440 445

Ser Ser Leu Asp Ala Pro Gln Thr Arg Gly His Asp Trp Arg Met Leu
450 455 460

Ala His Lys Leu Asn Leu Asp Arg Tyr Leu Asn Tyr Phe Ala Thr Lys
465 470 475 480

Ser Ser Pro Thr Gly Val Ile Leu Asp Leu Trp Glu Ala Gln Asn Phe
485 490 495

Pro Asp Gly Asn Leu Ser Met Leu Ala Ala Val Leu Glu Glu Met Gly
500 505 510

Arg His Glu Thr Val Val Ser Leu Ala Ala Glu Gly Gln Tyr
515 520 525

<210> 7
<211> 813
<212> DNA
<213> Homo sapiens

<400> 7
cagtgacctc tgtgtacaca ctgcttctgg ccctgaggac gtggccctct atgtgggcct 60
natcgccgtg gccgnntgcc tggtcctgct gctgcttgtc ctcatcctcg tttattgccg 120
gaagaaggag gggctggact canatgtggc tgactcgtcc attctcacct caggcttcca 180
gcccgtcagc atcaagccca gcaaagcaga caacccccat ctgctcacca tccagccgga 240
cctcagcacc accaccacca cctaccaggg cagtctctgt ccccggcagg atgggcccag 300
ccccaagttc cagctcacca atgggcacct gctcagcccc ctgggtggcg gccgccacac 360
actgcaccac agctctccca cctctgaggc cgaggagttc gtctcccgcc tctccaccca 420
gaactacttc cgctccctgc cccgaggcac cagcaacatg acctatggga ccttcaactt 480
cctcgggggc cggctgatga tccctaatac aggaatcaag cctcctcatc cccccagatg 540
ccatacccgg agggaagatc tatgagatct acctcacgct tgcacaagcc ggaagacgtg 600
aggttgcccc tagctggctg tcagaccctg ctgagtccca tcgttagctt gtggaccccc 660
tggcgtcctg ctcacccggc cagtcatcct ggctatggac cactgtgggg agcccagccc 720
tgacagctgg agcctgcgcc tcaaaaagca gtcgtgcgag ggcagctggg aggatgtgct 780
gcacctgggc gaggaggcgc cctcccactc act 813

<210> 8
<211> 265
<212> PRT
<213> Homo sapiens

<400> 8
Gln Pro Leu Cys Thr His Cys Phe Trp Pro Gly Arg Gly Pro Leu Cys
1 5 10 15

Gly Pro Xaa Arg Arg Gly Xaa Xaa Pro Gly Pro Ala Ala Ala Cys Pro
20 25 30

His Pro Arg Leu Leu Pro Glu Glu Gly Gly Ala Gly Leu Xaa Cys Gly
35 40 45

Leu Val His Ser His Leu Arg Leu Pro Ala Arg Gln His Gln Ala Gln
50 55 60

Gln Ser Arg Gln Pro Pro Ser Ala His His Pro Ala Gly Pro Gln His
65 70 75 80

His His His His Leu Pro Gly Gln Ser Leu Ser Pro Ala Gly Trp Ala
85 90 95

```
Gln Pro Gln Val Pro Ala His Gln Trp Ala Pro Ala Gln Pro Pro Gly
            100                 105                 110

Trp Arg Pro Pro His Thr Ala Pro Gln Leu Ser His Leu Gly Arg Gly
        115                 120                 125

Val Arg Leu Pro Pro Leu His Pro Glu Leu Leu Pro Leu Pro Ala Pro
    130                 135                 140

Arg His Gln Gln His Asp Leu Trp Asp Leu Gln Leu Pro Arg Gly Pro
145                 150                 155                 160

Ala Asp Asp Pro Tyr Arg Asn Gln Ala Ser Ser Ser Pro Gln Met Pro
                165                 170                 175

Tyr Pro Glu Gly Arg Ser Met Arg Ser Thr Ser Arg Leu His Lys Pro
            180                 185                 190

Glu Asp Val Arg Leu Pro Leu Ala Gly Cys Gln Thr Leu Leu Ser Pro
        195                 200                 205

Ile Val Ser Leu Trp Thr Pro Trp Arg Pro Ala His Pro Ala Ser His
    210                 215                 220

Pro Gly Tyr Gly Pro Leu Trp Gly Ala Gln Pro Gln Leu Glu Pro Ala
225                 230                 235                 240

Pro Gln Lys Ala Val Val Arg Gly Gln Leu Gly Gly Cys Ala Ala Pro
                245                 250                 255

Gly Arg Gly Gly Ala Leu Pro Leu Thr
            260                 265
```

<210> 9
<211> 266
<212> PRT
<213> Homo sapiens

<400> 9

```
Ser Asp Leu Cys Val His Thr Ala Ser Gly Pro Glu Asp Val Ala Leu
  1               5                  10                  15

Tyr Val Gly Xaa Ile Ala Val Ala Xaa Cys Leu Val Leu Leu Leu Leu
             20                  25                  30

Val Leu Ile Leu Val Tyr Cys Arg Lys Lys Glu Gly Leu Asp Ser Xaa
         35                  40                  45

Val Ala Asp Ser Ser Ile Leu Thr Ser Gly Phe Gln Pro Val Ser Ile
     50                  55                  60

Lys Pro Ser Lys Ala Asp Asn Pro His Leu Leu Thr Ile Gln Pro Asp
 65                  70                  75                  80

Leu Ser Thr Thr Thr Thr Thr Tyr Gln Gly Ser Leu Cys Pro Arg Gln
                 85                  90                  95
```

Asp Gly Pro Ser Pro Lys Phe Gln Leu Thr Asn Gly His Leu Leu Ser
            100                 105                 110

Pro Leu Gly Gly Gly Arg His Thr Leu His His Ser Ser Pro Thr Ser
        115                 120                 125

Glu Ala Glu Glu Phe Val Ser Arg Leu Ser Thr Gln Asn Tyr Phe Arg
    130                 135                 140

Ser Leu Pro Arg Gly Thr Ser Asn Met Thr Tyr Gly Thr Phe Asn Phe
145                 150                 155                 160

Leu Gly Gly Arg Leu Met Ile Pro Asn Thr Gly Ile Lys Pro Pro His
                165                 170                 175

Pro Pro Arg Cys His Thr Pro Arg Glu Asp Leu Asp Leu Pro His Ala
            180                 185                 190

Cys Thr Ser Arg Lys Thr Gly Cys Pro Leu Ala Val Arg Pro Cys Val
        195                 200                 205

Pro Ser Leu Ala Cys Gly Pro Pro Gly Val Leu Leu Thr Arg Pro Val
    210                 215                 220

Ile Leu Ala Met Asp His Cys Gly Glu Pro Ser Pro Asp Ser Trp Ser
225                 230                 235                 240

Leu Arg Leu Lys Lys Gln Ser Cys Glu Gly Ser Trp Glu Asp Val Leu
                245                 250                 255

His Leu Gly Glu Glu Ala Pro Ser His Ser
            260                 265

<210> 10
<211> 266
<212> PRT
<213> Homo sapiens

<400> 10
Val Thr Ser Val Tyr Thr Leu Leu Leu Ala Leu Arg Thr Trp Pro Ser
 1               5                  10                  15

Met Trp Ala Xaa Ser Pro Trp Pro Xaa Ala Trp Ser Cys Cys Cys Leu
            20                  25                  30

Ser Ser Ser Ser Phe Ile Ala Gly Arg Arg Arg Gly Trp Thr Xaa Met
        35                  40                  45

Trp Leu Thr Arg Pro Phe Ser Pro Gln Ala Ser Ser Pro Ser Ala Ser
    50                  55                  60

Ser Pro Ala Lys Gln Thr Thr Pro Ile Cys Ser Pro Ser Ser Arg Thr
65                  70                  75                  80

Ser Ala Pro Pro Pro Pro Pro Thr Arg Ala Val Ser Val Pro Gly Arg
                85                  90                  95

Met Gly Pro Ala Pro Ser Ser Ser Ser Pro Met Gly Thr Cys Ser Ala
100 105 110

Pro Trp Val Ala Ala Ala Thr His Cys Thr Thr Ala Leu Pro Pro Leu
115 120 125

Arg Pro Arg Ser Ser Ser Pro Ala Ser Pro Pro Arg Thr Thr Ser Ala
130 135 140

Pro Cys Pro Glu Ala Pro Ala Thr Pro Met Gly Pro Ser Thr Ser Ser
145 150 155 160

Gly Ala Gly Ser Leu Ile Gln Glu Ser Ser Leu Leu Ile Pro Pro Asp
165 170 175

Ala Ile Pro Arg Gly Lys Ile Tyr Glu Ile Tyr Leu Thr Leu Ala Gln
180 185 190

Ala Gly Arg Arg Glu Val Ala Pro Ser Trp Leu Ser Asp Pro Ala Glu
195 200 205

Ser His Arg Leu Val Asp Pro Leu Ala Ser Cys Ser Pro Gly Gln Ser
210 215 220

Ser Trp Leu Trp Thr Thr Val Gly Ser Pro Ala Leu Thr Ala Gly Ala
225 230 235 240

Cys Ala Ser Lys Ser Ser Arg Ala Arg Ala Ala Gly Arg Met Cys Cys
245 250 255

Thr Trp Ala Arg Arg Arg Pro Pro Thr His
260 265

<210> 11
<211> 6981
<212> DNA
<213> Caenorhabditis elegans

<400> 11
atgtcgtcgg tggatgacga gcatcatcct tcaattgaag ttgatattcc tgacgtttat 60
gaaagtgcat ctgactgtga aaaaactgaa gactgcacct atctaggatc tattctcaag 120
gcgaaaaaat cattgcgaaa aaccaagaaa gtcgctagtc ctggtatcta tccgggctcg 180
gcgggtgctc tttcatctct ttcaaatggc ctagccgatg agaatagatc gccaactagt 240
gctgaagttt ctgattcagt atgtggacaa aaatctataa attcagttga cttgaggttc 300
agaggattga gagatgaacg tgaactcgtg caaaagaaaa catttacaaa atgggtcaac 360
tcacacttgg ttcgagtcag ctgcaaagta caagacctct acatggatat gcgagacgga 420
aaaatgctcc tccgactcct cgcagttctt tccggtgaac gtctgccgaa accgacgccg 480
ggaaaaatgc gtatccactg tcttgagaac gttgagaaag gtttgcaatt cttgcgcaat 540
caacacgtgc atctcgaaaa cttgggttca cacgatatcg ttgacggaaa ttcacgcctc 600
acgttgggtc tcatctggac gatcatcctt cgattccaga tccaagatat cacatttgaa 660
gacgctgaca atcacgagac acgttcagcg aaggaggcct tgttactgtg gtgccagatg 720
aaaactgctg gatatccaaa tgtcaacgtc aagaacttct caaccagttg gagagatggt 780
cttgcattca atgcacttat tcataaacat agacctgatc tcgtagacta cgataattta 840
caaaaatcga atgctcttta caatttgcaa tctgcttttg ataccgcaga aaatcaattg 900
ggattggcca agttcctcga tgccgaagat gtgaacgttg accaacctga tgagaaatcc 960
attatcactt atgtagttac atattatcat tacttcaaca aactgaaaca agataacatc 1020

caaggaaaac gtatcggaaa agtgattaat gagttgatgg agaacgacaa aatgatcaat 1080
cgatacgaaa ctctttcatc tgatttgctc gaatggatca atgccaagat ccaacttctc 1140
aacgaacgtc attttgagaa taatttggaa ggtgttcaaa gacagctcac cgagttcaac 1200
gactaccgta ctcaggagaa accaccaaag ttcgatgaaa agggagagct cgaagtactt 1260
ctgttcactc ttcaatcagc tatgcgtgcc aacaatcaga gacctttcgt gccgagagaa 1320
ggaaaactga ttgctgatat caaccgagca tggcaatccc ttgagaaggc agaacacgaa 1380
cgagaacttg tcctgaaaga ggaactcatt cgtcaagaga aactcgaaca actcgcagcc 1440
cgattcaaca gaaaggccga gatgcgagag acttggctca cagaaaacca acggcttgtc 1500
agtcaggaca atttcggaaa cgacttgtct tcggtggaag ctgccaccaa aaagcacgaa 1560
gccatcgaga cggacatctt tgcttacgag gagcgtgttc aagccgtcgt cgctgttgct 1620
ggagagcttg aagccgagaa ctaccatgat caagccaaaa tcaatgagag aaaggagaat 1680
gttcttcaac tctggaacta tcttttccaa cttcttttgg ctcgcagagt tcgattggag 1740
ttgtcaatgg ccattcaaaa gatcttccac gatatgcttc ttactctcga tctcatggat 1800
gatattaaga gcagactttt aagcgaggat cttggagccc atcttatgga cgtagaagat 1860
ttgcttcaaa agcacgctct tcttgagtct gatattaata ttattggaga acgagttaac 1920
aattctattg ctcaagctca acgcttccga aatccagatg gaccagatgg aagtggatac 1980
aagccagttg aaccaggaac cattgatgag agaagtgatg ttctacaaaa acgttacaag 2040
gagcttcttg atttggctgc tgaacgtaag cgtcgtctcg aagacaacaa gagactttgc 2100
cagttctggt gggatgttgc tgagcttgaa catggtatta aggaacaaga gcaagttctt 2160
tcgtccactg atactggccg tgatattgtc actgtttcgc atttgttggc caaacataag 2220
aacgctgaga acaaccttcg cgacttggag aagtacctcg atcgtcttga tgtttctgga 2280
aaagaattgc aagatgaaag catcccagga agtgacaaca ttcctccaag gcttgccgag 2340
attcgtgatt acatcaacaa gctcaaggaa ctctctgctt ctcgcaaaga aagacttgcc 2400
ggtggagtcg aatactacca attcttcacc gacgccgacg atgtcgatcg ctacttgtat 2460
gatactctcc gcgtcatgtc atccgaagac gttggaaaag acgagggaac tgttcaactt 2520
ctcctcaaga agcacgacga tgtccatgac gagctccaga actttgacca acacatcaag 2580
gttcttcacg ccaaggccga gtctcttcca caggaagcca gagaacatcc agatatcaga 2640
caacgattgg acacaactct caagcagaag gccgaactcg aaaacctttc tcaacttcgc 2700
aagcagagac ttatcgatgc tctctccctg tacaagttgt actcggatgc cgattcagtc 2760
gagtcttgga tcgatgagaa gggaaaactt ttggcaacat tggtccctgg acgcgatatc 2820
gaagaggttg agatcatgaa gcatcgtttc gatactcttg agcaagatat gaagaatcaa 2880
gaagctaagg ttactaacgt caatgacttg gctcgtcaat tgcttaatgt ggaacatcca 2940
aactccgacg acattcttca ccgtcaaaac aagctcaacg ctcgctgggc tcaacttcgc 3000
gacatggtcg atcaaaagag aaacgagctc gagagagctc accgcctcga aactttccgc 3060
atcgattgcc aagaaacagt tacttggatc gaggacaaga ctcgtgtact cgaggattcg 3120
gacgcgttga ccaacgatct ttccggtgtc atgaaacttc aaagaagatt gtccatgatg 3180
gaacgcgatt tgggagcgat tcaagccaag ctcgactcgc ttcacaaaga agccgatgac 3240
atcgaaagag aacgaccaca agaagctcaa gccattcgtg aagatatcaa gagaatccac 3300
caagtctggg atatccttaa caagaaagtt cgtgagcacg aagctaaact cgatgaagct 3360
ggagatcttc aacgcttcct ccgtgatttg gatcacttcc aagcctggct tactgccacc 3420
cagagacaag ttgcctcaga ggaagagcca caatccttgg ctgaagctga gcaacttctc 3480
aatcaacacg ctgccatccg tgaagagatt gacggatacg ctgaggatta caagaagatg 3540
cgtgcaatgg gagatcgtgt cactcaagat caaaccgatc cacaatacat gttccttcgt 3600
caacgtcttg ctggacttca agaaggatgg gaagagttgc agcgaatgtg ggacaacaga 3660
caacatttgc tttctcaagg actcaacctt caaatgttcc ttcgcgatgc caaacaagca 3720
gaggttatgc tctctcaaca agaaaactat ctcgccaagg acgatatccc acaatctttg 3780
gagcaagccg agaaccaact caagagacac caagacttca tcaccactat ggatgccaac 3840
gacgaaaaga ttcgcgccgt cggaatgttc ggtgatcagc tctgccaaga tggtcactat 3900
gccgctgaca agatccataa gaaggctaga aacattgatg aacgtcgtgg agccaaccgc 3960
gagaaggctc aagaagtact caagaagctg aaggacgcct tgtctctcca acaattcttg 4020
tccgactgcg acgagctgag agaatggatc gaggagaaga tgatccgtgc tcaagatgag 4080
acttacagag acgccaagac tatcacctcc aaattcgtcc gacatcaagc attccaatcc 4140
gaacttgctg ctaacaagga acgtcttgac caacttaaac atgctgccat caaccttggt 4200
gacgacaagc cagagtacca cggaaccatt gatccacaga tcgaggaact cgctactcaa 4260
tgggatgagt tggaaaagac tacagaggag aagggacaga agttgtttga tgctaaccgc 4320
caacaacttt acgtacaaag tattgctgac atgaaggaat gggcaaccca acttgagaac 4380
gagatgactc gtgaagacca accaggagat cttaccactg tcaatgtggc catgcaaaaa 4440
caacatttga tcgaaactga aatgatcaag aaggctcaac acattgatca actcatggag 4500

atggagccac agctcgagga acttcaccca gacgagcttg aaaacattaa ggctcatcgt 4560
ttggcagttc aagaacaact tcaaagactc caggctccat tggacgaccg acgtaaggca 4620
ttggaacgca agaaggccgc cttccaattt ggtcgcgatg ttgatgacga gaaattgtgg 4680
atttctgaga gacttgttct cgctaaagct caaaatcttg gagaaagcct tccagattgt 4740
catcgtctcc agaagaacct acaacttctt tccaatgaaa tcgacaacca cgaaccatgg 4800
atcaaccaga tctgtaataa tggtcaagag ttgattgacg aaggacatgc taacggacca 4860
gcattcgaga agaagattca agaactgcgc agtgcttggc aagagttgaa ggaagccgtc 4920
aaagacagaa agggagatct tggagaaagc gagaaggctc atcaattctt gtatgattgc 4980
ggagaggctg aagcctggat gagcgaacaa gaattgtaca tgatgcaaga cgaaagagga 5040
aaagacgagt tctcaaccaa gaaccaaatc aagaaacacg aacgccttca atccgatatt 5100
gacaaatttg ctgatacgat ccgtgctctt gccaccaagg ctcacaagtt cgtcgaagag 5160
aagtcaccac tcaccgaaca gatccaagtt cgtcaagctc aaatcgagaa gctctacgca 5220
ggtcttcaag atttgtcgaa ggagcgcaga aagagactcg aagagacctt ggaactttac 5280
gctcttcatc gtgaaatcga cgatcttctt caatggatcg ccgacaagga agttgttgca 5340
ggaagtcaag agaacggaca agactacgaa cacgtccaaa tgcttcaaga gagattccaa 5400
cagtttgctc gtgacaccga aaacatcgga tccgaacgtg ttgccaatgc taatgatgga 5460
tgtgatactc tcattggtca cggacacact gacgctccga ctatagcatt gtggaaggat 5520
tcgttgaatg aggcctggga gaatcttctg gagttgatgg acactcgtgc ccaaattctc 5580
gaggcttcta gacttcttca caagttctac cacgactgcc gtgattgctt gtcaagaatc 5640
atggaaaaga ctcatgctat gcctgatgat cttggacgtg actcaagcag tgttggtgca 5700
ctctccagaa aacatcagaa ctatctcaag gatattgctg caattggaga acaagttgct 5760
caaatcgaaa gagatgccgc agagcttcga gatggttacg ctggagacaa ggctcttgat 5820
atcggatctc gcgaaagcga ggttgtcaag gcatggaggc atctccgcgg actctgtgac 5880
gcccgtacat ctcgtctcat ggatacttct gatcttttca agttcatgaa tatggtccgt 5940
gatttgttgt tgtggatgga tgaggtgaag agagaaatga actcacaaga acgtccgaag 6000
gatgtgagcg gagtcgagtt gctcatgaac aatcatcaat cgctcaaggc tgaaatcgat 6060
gctcgcgagg agaacttcaa cgcgtgcatc tcattgggac gagatctgct caacagaaaa 6120
cactacgcat caagtgaaat cgagaagaag ctgatcaagt tgacaaccga aagagccgaa 6180
atgatgcgca gatgggaaga tcgttgggaa tatcttcagc tcattctgga agtgtaccaa 6240
ttcgcccgcg acgctgctgt cgccgaatct tggctcttcg cccaagaacc atatcttatt 6300
tcgaaggaat acggtagaaa cctcgaagaa accatcaaac tcatcaagaa acacgaggcg 6360
ttcgagaaat cggcattcgc acaagaagaa cgtttcttgg cgctcgagaa attgacaacg 6420
ttcgaattga aagaaactca acaccgcgag gaggagaccg ccaagcgacg tggtccagct 6480
catattggaa gtccttctag gtcaactcca gccgcggaaa catctttcgg cgcccaggat 6540
gatggagcca aacaaggaga agcattcgaa ggaacgctta tcagaaagca cacttacgag 6600
agtctcgacc ggaaagcggc taaccgctca tgggagaaac tttacgcagt gttgcggcag 6660
aatgaactgt cattctacaa agatccgaag catagagatg aatctgtgca tggagagcca 6720
ccaatggcac tgccaggatg ctctgtcaac gtggcatcag actatcaaaa gaagaagaac 6780
gtactgtcgc tgaggttacc catcggcgcc gagtacctat tccaatgcgg tagcgaagaa 6840
gacatgcagc gatggctcac cgagttgcaa gtcgccacgg ggcaagccca actggaagag 6900
gctagtagat cacagacgct accagcagaa ggcagtgcca caaagaagaa gggaggattt 6960
ttctctcgag gcaagaagta a 6981

<210> 12
<211> 2326
<212> PRT
<213> Caenorhabditis elegans

<400> 12
Met Ser Ser Val Asp Asp Glu His His Pro Ser Ile Glu Val Asp Ile
1 5 10 15

Pro Asp Val Tyr Glu Ser Ala Ser Asp Cys Glu Lys Thr Glu Asp Cys
20 25 30

Thr Tyr Leu Gly Ser Ile Leu Lys Ala Lys Lys Ser Leu Arg Lys Thr
35 40 45

```
Lys Lys Val Ala Ser Pro Gly Ile Tyr Pro Gly Ser Ala Gly Ala Leu
    50                  55                  60

Ser Ser Leu Ser Asn Gly Leu Ala Asp Glu Asn Arg Ser Pro Thr Ser
65                  70                  75                  80

Ala Glu Val Ser Asp Ser Val Cys Gly Gln Lys Ser Ile Asn Ser Val
                85                  90                  95

Asp Leu Arg Phe Arg Gly Leu Arg Asp Glu Arg Glu Leu Val Gln Lys
            100                 105                 110

Lys Thr Phe Thr Lys Trp Val Asn Ser His Leu Val Arg Val Ser Cys
        115                 120                 125

Lys Val Gln Asp Leu Tyr Met Asp Met Arg Asp Gly Lys Met Leu Leu
    130                 135                 140

Arg Leu Leu Ala Val Leu Ser Gly Glu Arg Leu Pro Lys Pro Thr Pro
145                 150                 155                 160

Gly Lys Met Arg Ile His Cys Leu Glu Asn Val Glu Lys Gly Leu Gln
                165                 170                 175

Phe Leu Arg Asn Gln His Val His Leu Glu Asn Leu Gly Ser His Asp
            180                 185                 190

Ile Val Asp Gly Asn Ser Arg Leu Thr Leu Gly Leu Ile Trp Thr Ile
        195                 200                 205

Ile Leu Arg Phe Gln Ile Gln Asp Ile Thr Phe Glu Asp Ala Asp Asn
    210                 215                 220

His Glu Thr Arg Ser Ala Lys Glu Ala Leu Leu Leu Trp Cys Gln Met
225                 230                 235                 240

Lys Thr Ala Gly Tyr Pro Asn Val Asn Val Lys Asn Phe Ser Thr Ser
                245                 250                 255

Trp Arg Asp Gly Leu Ala Phe Asn Ala Leu Ile His Lys His Arg Pro
            260                 265                 270

Asp Leu Val Asp Tyr Asp Asn Leu Gln Lys Ser Asn Ala Leu Tyr Asn
        275                 280                 285

Leu Gln Ser Ala Phe Asp Thr Ala Glu Asn Gln Leu Gly Leu Ala Lys
    290                 295                 300

Phe Leu Asp Ala Glu Asp Val Asn Val Asp Gln Pro Asp Glu Lys Ser
305                 310                 315                 320

Ile Ile Thr Tyr Val Val Thr Tyr Tyr His Tyr Phe Asn Lys Leu Lys
                325                 330                 335

Gln Asp Asn Ile Gln Gly Lys Arg Ile Gly Lys Val Ile Asn Glu Leu
            340                 345                 350

Met Glu Asn Asp Lys Met Ile Asn Arg Tyr Glu Thr Leu Ser Ser Asp
```

355 360 365

Leu Leu Glu Trp Ile Asn Ala Lys Ile Gln Leu Leu Asn Glu Arg His
370 375 380

Phe Glu Asn Asn Leu Glu Gly Val Gln Arg Gln Leu Thr Glu Phe Asn
385 390 395 400

Asp Tyr Arg Thr Gln Glu Lys Pro Pro Lys Phe Asp Glu Lys Gly Glu
405 410 415

Leu Glu Val Leu Leu Phe Thr Leu Gln Ser Ala Met Arg Ala Asn Asn
420 425 430

Gln Arg Pro Phe Val Pro Arg Glu Gly Lys Leu Ile Ala Asp Ile Asn
435 440 445

Arg Ala Trp Gln Ser Leu Glu Lys Ala Glu His Glu Arg Glu Leu Val
450 455 460

Leu Lys Glu Glu Leu Ile Arg Gln Glu Lys Leu Glu Gln Leu Ala Ala
465 470 475 480

Arg Phe Asn Arg Lys Ala Glu Met Arg Glu Thr Trp Leu Thr Glu Asn
485 490 495

Gln Arg Leu Val Ser Gln Asp Asn Phe Gly Asn Asp Leu Ser Ser Val
500 505 510

Glu Ala Ala Thr Lys Lys His Glu Ala Ile Glu Thr Asp Ile Phe Ala
515 520 525

Tyr Glu Glu Arg Val Gln Ala Val Val Ala Val Ala Gly Glu Leu Glu
530 535 540

Ala Glu Asn Tyr His Asp Gln Ala Lys Ile Asn Glu Arg Lys Glu Asn
545 550 555 560

Val Leu Gln Leu Trp Asn Tyr Leu Phe Gln Leu Leu Leu Ala Arg Arg
565 570 575

Val Arg Leu Glu Leu Ser Met Ala Ile Gln Lys Ile Phe His Asp Met
580 585 590

Leu Leu Thr Leu Asp Leu Met Asp Asp Ile Lys Ser Arg Leu Leu Ser
595 600 605

Glu Asp Leu Gly Ala His Leu Met Asp Val Glu Asp Leu Leu Gln Lys
610 615 620

His Ala Leu Leu Glu Ser Asp Ile Asn Ile Ile Gly Glu Arg Val Asn
625 630 635 640

Asn Ser Ile Ala Gln Ala Gln Arg Phe Arg Asn Pro Asp Gly Pro Asp
645 650 655

Gly Ser Gly Tyr Lys Pro Val Glu Pro Gly Thr Ile Asp Glu Arg Ser
660 665 670

Asp Val Leu Gln Lys Arg Tyr Lys Glu Leu Leu Asp Leu Ala Ala Glu
675 680 685

Arg Lys Arg Arg Leu Glu Asp Asn Lys Arg Leu Cys Gln Phe Trp Trp
690 695 700

Asp Val Ala Glu Leu Glu His Gly Ile Lys Glu Gln Glu Gln Val Leu
705 710 715 720

Ser Ser Thr Asp Thr Gly Arg Asp Ile Val Thr Val Ser His Leu Leu
725 730 735

Ala Lys His Lys Asn Ala Glu Asn Asn Leu Arg Asp Leu Glu Lys Tyr
740 745 750

Leu Asp Arg Leu Asp Val Ser Gly Lys Glu Leu Gln Asp Glu Ser Ile
755 760 765

Pro Gly Ser Asp Asn Ile Pro Pro Arg Leu Ala Glu Ile Arg Asp Tyr
770 775 780

Ile Asn Lys Leu Lys Glu Leu Ser Ala Ser Arg Lys Glu Arg Leu Ala
785 790 795 800

Gly Gly Val Glu Tyr Tyr Gln Phe Phe Thr Asp Ala Asp Asp Val Asp
805 810 815

Arg Tyr Leu Tyr Asp Thr Leu Arg Val Met Ser Ser Glu Asp Val Gly
820 825 830

Lys Asp Glu Gly Thr Val Gln Leu Leu Leu Lys Lys His Asp Asp Val
835 840 845

His Asp Glu Leu Gln Asn Phe Asp Gln His Ile Lys Val Leu His Ala
850 855 860

Lys Ala Glu Ser Leu Pro Gln Glu Ala Arg Glu His Pro Asp Ile Arg
865 870 875 880

Gln Arg Leu Asp Thr Thr Leu Lys Gln Lys Ala Glu Leu Glu Asn Leu
885 890 895

Ser Gln Leu Arg Lys Gln Arg Leu Ile Asp Ala Leu Ser Leu Tyr Lys
900 905 910

Leu Tyr Ser Asp Ala Asp Ser Val Glu Ser Trp Ile Asp Glu Lys Gly
915 920 925

Lys Leu Leu Ala Thr Leu Val Pro Gly Arg Asp Ile Glu Glu Val Glu
930 935 940

Ile Met Lys His Arg Phe Asp Thr Leu Glu Gln Asp Met Lys Asn Gln
945 950 955 960

Glu Ala Lys Val Thr Asn Val Asn Asp Leu Ala Arg Gln Leu Leu Asn
965 970 975

Val Glu His Pro Asn Ser Asp Asp Ile Leu His Arg Gln Asn Lys Leu
980 985 990

Asn Ala Arg Trp Ala Gln Leu Arg Asp Met Val Asp Gln Lys Arg Asn
995 1000 1005

Glu Leu Glu Arg Ala His Arg Leu Glu Thr Phe Arg Ile Asp Cys Gln
1010 1015 1020

Glu Thr Val Thr Trp Ile Glu Asp Lys Thr Arg Val Leu Glu Asp Ser
1025 1030 1035 1040

Asp Ala Leu Thr Asn Asp Leu Ser Gly Val Met Lys Leu Gln Arg Arg
1045 1050 1055

Leu Ser Met Met Glu Arg Asp Leu Gly Ala Ile Gln Ala Lys Leu Asp
1060 1065 1070

Ser Leu His Lys Glu Ala Asp Asp Ile Glu Arg Glu Arg Pro Gln Glu
1075 1080 1085

Ala Gln Ala Ile Arg Glu Asp Ile Lys Arg Ile His Gln Val Trp Asp
1090 1095 1100

Ile Leu Asn Lys Lys Val Arg Glu His Glu Ala Lys Leu Asp Glu Ala
1105 1110 1115 1120

Gly Asp Leu Gln Arg Phe Leu Arg Asp Leu Asp His Phe Gln Ala Trp
1125 1130 1135

Leu Thr Ala Thr Gln Arg Gln Val Ala Ser Glu Glu Glu Pro Gln Ser
1140 1145 1150

Leu Ala Glu Ala Glu Gln Leu Leu Asn Gln His Ala Ala Ile Arg Glu
1155 1160 1165

Glu Ile Asp Gly Tyr Ala Glu Asp Tyr Lys Lys Met Arg Ala Met Gly
1170 1175 1180

Asp Arg Val Thr Gln Asp Gln Thr Asp Pro Gln Tyr Met Phe Leu Arg
1185 1190 1195 1200

Gln Arg Leu Ala Gly Leu Gln Glu Gly Trp Glu Glu Leu Gln Arg Met
1205 1210 1215

Trp Asp Asn Arg Gln His Leu Leu Ser Gln Gly Leu Asn Leu Gln Met
1220 1225 1230

Phe Leu Arg Asp Ala Lys Gln Ala Glu Val Met Leu Ser Gln Gln Glu
1235 1240 1245

Asn Tyr Leu Ala Lys Asp Asp Ile Pro Gln Ser Leu Glu Gln Ala Glu
1250 1255 1260

Asn Gln Leu Lys Arg His Gln Asp Phe Ile Thr Thr Met Asp Ala Asn
1265 1270 1275 1280

Asp Glu Lys Ile Arg Ala Val Gly Met Phe Gly Asp Gln Leu Cys Gln

1285 1290 1295

Asp Gly His Tyr Ala Ala Asp Lys Ile His Lys Lys Ala Arg Asn Ile
1300 1305 1310

Asp Glu Arg Arg Gly Ala Asn Arg Glu Lys Ala Gln Glu Val Leu Lys
1315 1320 1325

Lys Leu Lys Asp Ala Leu Ser Leu Gln Gln Phe Leu Ser Asp Cys Asp
1330 1335 1340

Glu Leu Arg Glu Trp Ile Glu Glu Lys Met Ile Arg Ala Gln Asp Glu
1345 1350 1355 1360

Thr Tyr Arg Asp Ala Lys Thr Ile Thr Ser Lys Phe Val Arg His Gln
1365 1370 1375

Ala Phe Gln Ser Glu Leu Ala Ala Asn Lys Glu Arg Leu Asp Gln Leu
1380 1385 1390

Lys His Ala Ala Ile Asn Leu Gly Asp Asp Lys Pro Glu Tyr His Gly
1395 1400 1405

Thr Ile Asp Pro Gln Ile Glu Glu Leu Ala Thr Gln Trp Asp Glu Leu
1410 1415 1420

Glu Lys Thr Thr Glu Glu Lys Gly Gln Lys Leu Phe Asp Ala Asn Arg
1425 1430 1435 1440

Gln Gln Leu Tyr Val Gln Ser Ile Ala Asp Met Lys Glu Trp Ala Thr
1445 1450 1455

Gln Leu Glu Asn Glu Met Thr Arg Glu Asp Gln Pro Gly Asp Leu Thr
1460 1465 1470

Thr Val Asn Val Ala Met Gln Lys Gln His Leu Ile Glu Thr Glu Met
1475 1480 1485

Ile Lys Lys Ala Gln His Ile Asp Gln Leu Met Glu Met Glu Pro Gln
1490 1495 1500

Leu Glu Glu Leu His Pro Asp Glu Leu Glu Asn Ile Lys Ala His Arg
1505 1510 1515 1520

Leu Ala Val Gln Glu Gln Leu Gln Arg Leu Gln Ala Pro Leu Asp Asp
1525 1530 1535

Arg Arg Lys Ala Leu Glu Arg Lys Lys Ala Ala Phe Gln Phe Gly Arg
1540 1545 1550

Asp Val Asp Asp Glu Lys Leu Trp Ile Ser Glu Arg Leu Val Leu Ala
1555 1560 1565

Lys Ala Gln Asn Leu Gly Glu Ser Leu Pro Asp Cys His Arg Leu Gln
1570 1575 1580

Lys Asn Leu Gln Leu Leu Ser Asn Glu Ile Asp Asn His Glu Pro Trp
1585 1590 1595 1600

Ile Asn Gln Ile Cys Asn Asn Gly Gln Glu Leu Ile Asp Glu Gly His
1605 1610 1615

Ala Asn Gly Pro Ala Phe Glu Lys Lys Ile Gln Glu Leu Arg Ser Ala
1620 1625 1630

Trp Gln Glu Leu Lys Glu Ala Val Lys Asp Arg Lys Gly Asp Leu Gly
1635 1640 1645

Glu Ser Glu Lys Ala His Gln Phe Leu Tyr Asp Cys Gly Glu Ala Glu
1650 1655 1660

Ala Trp Met Ser Glu Gln Glu Leu Tyr Met Met Gln Asp Glu Arg Gly
1665 1670 1675 1680

Lys Asp Glu Phe Ser Thr Lys Asn Gln Ile Lys Lys His Glu Arg Leu
1685 1690 1695

Gln Ser Asp Ile Asp Lys Phe Ala Asp Thr Ile Arg Ala Leu Ala Thr
1700 1705 1710

Lys Ala His Lys Phe Val Glu Glu Lys Ser Pro Leu Thr Glu Gln Ile
1715 1720 1725

Gln Val Arg Gln Ala Gln Ile Glu Lys Leu Tyr Ala Gly Leu Gln Asp
1730 1735 1740

Leu Ser Lys Glu Arg Arg Lys Arg Leu Glu Glu Thr Leu Glu Leu Tyr
1745 1750 1755 1760

Ala Leu His Arg Glu Ile Asp Asp Leu Leu Gln Trp Ile Ala Asp Lys
1765 1770 1775

Glu Val Val Ala Gly Ser Gln Glu Asn Gly Gln Asp Tyr Glu His Val
1780 1785 1790

Gln Met Leu Gln Glu Arg Phe Gln Gln Phe Ala Arg Asp Thr Glu Asn
1795 1800 1805

Ile Gly Ser Glu Arg Val Ala Asn Ala Asn Asp Gly Cys Asp Thr Leu
1810 1815 1820

Ile Gly His Gly His Thr Asp Ala Pro Thr Ile Ala Leu Trp Lys Asp
1825 1830 1835 1840

Ser Leu Asn Glu Ala Trp Glu Asn Leu Leu Glu Leu Met Asp Thr Arg
1845 1850 1855

Ala Gln Ile Leu Glu Ala Ser Arg Leu Leu His Lys Phe Tyr His Asp
1860 1865 1870

Cys Arg Asp Cys Leu Ser Arg Ile Met Glu Lys Thr His Ala Met Pro
1875 1880 1885

Asp Asp Leu Gly Arg Asp Ser Ser Ser Val Gly Ala Leu Ser Arg Lys
1890 1895 1900

His Gln Asn Tyr Leu Lys Asp Ile Ala Ala Ile Gly Glu Gln Val Ala
1905 1910 1915 1920

Gln Ile Glu Arg Asp Ala Ala Glu Leu Arg Asp Gly Tyr Ala Gly Asp
1925 1930 1935

Lys Ala Leu Asp Ile Gly Ser Arg Glu Ser Glu Val Val Lys Ala Trp
1940 1945 1950

Arg His Leu Arg Gly Leu Cys Asp Ala Arg Thr Ser Arg Leu Met Asp
1955 1960 1965

Thr Ser Asp Leu Phe Lys Phe Met Asn Met Val Arg Asp Leu Leu Leu
1970 1975 1980

Trp Met Asp Glu Val Lys Arg Glu Met Asn Ser Gln Glu Arg Pro Lys
1985 1990 1995 2000

Asp Val Ser Gly Val Glu Leu Leu Met Asn Asn His Gln Ser Leu Lys
2005 2010 2015

Ala Glu Ile Asp Ala Arg Glu Glu Asn Phe Asn Ala Cys Ile Ser Leu
2020 2025 2030

Gly Arg Asp Leu Leu Asn Arg Lys His Tyr Ala Ser Ser Glu Ile Glu
2035 2040 2045

Lys Lys Leu Ile Lys Leu Thr Thr Glu Arg Ala Glu Met Met Arg Arg
2050 2055 2060

Trp Glu Asp Arg Trp Glu Tyr Leu Gln Leu Ile Leu Glu Val Tyr Gln
2065 2070 2075 2080

Phe Ala Arg Asp Ala Ala Val Ala Glu Ser Trp Leu Phe Ala Gln Glu
2085 2090 2095

Pro Tyr Leu Ile Ser Lys Glu Tyr Gly Arg Asn Leu Glu Glu Thr Ile
2100 2105 2110

Lys Leu Ile Lys Lys His Glu Ala Phe Glu Lys Ser Ala Phe Ala Gln
2115 2120 2125

Glu Glu Arg Phe Leu Ala Leu Glu Lys Leu Thr Thr Phe Glu Leu Lys
2130 2135 2140

Glu Thr Gln His Arg Glu Glu Glu Thr Ala Lys Arg Arg Gly Pro Ala
2145 2150 2155 2160

His Ile Gly Ser Pro Ser Arg Ser Thr Pro Ala Ala Glu Thr Ser Phe
2165 2170 2175

Gly Ala Gln Asp Asp Gly Ala Lys Gln Gly Glu Ala Phe Glu Gly Thr
2180 2185 2190

Leu Ile Arg Lys His Thr Tyr Glu Ser Leu Asp Arg Lys Ala Ala Asn
2195 2200 2205

Arg Ser Trp Glu Lys Leu Tyr Ala Val Leu Arg Gln Asn Glu Leu Ser

2210 2215 2220

Phe Tyr Lys Asp Pro Lys His Arg Asp Glu Ser Val His Gly Glu Pro
2225 2230 2235 2240

Pro Met Ala Leu Pro Gly Cys Ser Val Asn Val Ala Ser Asp Tyr Gln
2245 2250 2255

[illegible] Lys Lys Asn Val Leu Ser Leu Arg Leu Pro Ile Gly Ala Glu Tyr
2260 2265 2270

[illegible] Phe Gln Cys Gly Ser Glu Glu Asp Met Gln Arg Trp Leu Thr Glu
2275 2280 2285

[illegible] [illegible] Val Ala Thr Gly Gln Ala Gln Leu Glu Glu Ala Ser Arg Ser
2290 2295 2300

[illegible] Thr Leu Pro Ala Glu Gly Ser Ala Thr Lys Lys Lys Gly Gly Phe
[illegible] 2310 2315 2320

[illegible] Ser Arg Gly Lys Lys
2325

<210> 13
<211> 2192
<212> DNA
<213> Caenorhabditis elegans

<400> 13
gaacttgctg ctaacaagga acgtcttgac caacttaaac atgctgccat caaccttggt 60
gacgacaagc cagagtacca cggaaccatt gatccacaga tcgaggaact cgctactcaa 120
tgggatgagt tggaaaagac tacagaggag aagggacaga agttgtttga tgctaaccgc 180
caacaacttt acgtacaaag tattgctgac atgaaggaat gggcaaccca acttgagaac 240
gagatgactc gtgaagacca accaggagat cttaccactg tcaatgtggc catgcaaaaa 300
caacatttga tcgaaactga aatgatcaag aaggctcaac acattgatca actcatggag 360
atggagccac agctcgagga acttcaccca gacgagcttg aaaacattaa ggctcatcgt 420
ttggcagttc aagaacaact tcaaagactc caggctccat tggacgaccg acgtaaggca 480
ttggaacgca agaaggccgc cttccaattt ggtcgcgatg ttgatgacga gaaattgtgg 540
atttctgaga gacttgttct cgctaaagct caaaatcttg gagaaagcct tccagattgt 600
catcgtctcc agaagaacct acaacttctt tccaatgaaa tcgacaacca cgaaccatgg 660
atcaaccaga tctgtaataa tggtcaagag ttgattgacg aaggacatgc taacggacca 720
gcattcgaga agaagattca agaactgcgc agtgcttggc aagagttgaa ggaagccgtc 780
aaagacagaa agggagatct tggagaaagc gagaaggctc atcaattctt gtatgattgc 840
ggagaggctg aagcctggat gagcgaacaa gaattgtaca tgatgcaaga cgaaagagga 900
aaagacgagt tctcaaccaa gaaccaaatc aagaaacacg aacgccttca atccgatatt 960
gacaaatttg ctgatacgat ccgtgctctt gccaccaagg ctcacaagtt cgtcgaagag 1020
aagtcaccac tcaccgaaca gatccaagtt cgtcaagctc aaatcgagaa gctctacgca 1080
ggtcttcaag atttgtcgaa ggagcgcaga aagagactcg aagagacctt ggaactttac 1140
gctcttcatc gtgaaatcga cgatcttctt caatggatcg ccgacaagga agttgttgca 1200
ggaagtcaag agaacggaca agactacgaa cacgtccaaa tgcttcaaga gagattccaa 1260
cagtttgctc gtgacaccga aaacatcgga tccgaacgtg ttgccaatgc taatgatgga 1320
tgtgatactc tcattggtca cggacacact gacgctccga ctatagcatt gtggaaggat 1380
tcgttgaatg aggcctggga gaatcttctg gagttgatgg acactcgtgc ccaaattctc 1440
gaggcttcta gacttcttca caagttctac cacgactgcc gtgattgctt gtcaagaatc 1500
atggaaaaga ctcatgctat gcctgatgat cttggacgtg actcaagcag tgttggtgca 1560
ctctccagaa aacatcagaa ctatctcaag gatattgctg caattggaga acaagttgct 1620
caaatcgaaa gagatgccgc agagcttcga gatggttacg ctggagacaa ggctcttgat 1680

```
atcggatctc gcgaaagcga ggttgtcaag gcatggaggc atctccgcgg actctgtgac 1740
gcccgtacat ctcgtctcat ggatacttct gatcttttca agttcatgaa tatggtccgt 1800
gatttgttgt tgtggatgga tgaggtgaag agagaaatga actcacaaga acgtccgaag 1860
gatgtgagcg gagtcgagtt gctcatgaac aatcatcaat cgctcaaggc tgaaatcgat 1920
gctcgcgagg agaacttcaa cgcgtgcatc tcattgggac gagatctgct caacagaaaa 1980
cactacgcat caagtgaaat cgagaagaag ctgatcaagt tgacaaccga aagagccgaa 2040
atgatgcgca gatgggaaga tcgttgggaa tatcttcagc tcattctgga agtgtaccaa 2100
ttcgcccgcg acgctgctgt cgccgaatct tggctcttcg cccaagaacc atatcttatt 2160
tcgaaggaat acggtagaaa cctcgaagaa ac                               2192
```

<210> 14
<211> 731
<212> PRT
<213> Caenorhabditis elegans

<400> 14

```
Glu Leu Ala Ala Asn Lys Glu Arg Leu Asp Gln Leu Lys His Ala Ala
 1               5                  10                  15

Ile Asn Leu Gly Asp Asp Lys Pro Glu Tyr His Gly Thr Ile Asp Pro
            20                  25                  30

Gln Ile Glu Glu Leu Ala Thr Gln Trp Asp Glu Leu Glu Lys Thr Thr
        35                  40                  45

Glu Glu Lys Gly Gln Lys Leu Phe Asp Ala Asn Arg Gln Gln Leu Tyr
    50                  55                  60

Val Gln Ser Ile Ala Asp Met Lys Glu Trp Ala Thr Gln Leu Glu Asn
65                  70                  75                  80

Glu Met Thr Arg Glu Asp Gln Pro Gly Asp Leu Thr Thr Val Asn Val
                85                  90                  95

Ala Met Gln Lys Gln His Leu Ile Glu Thr Glu Met Ile Lys Lys Ala
            100                 105                 110

Gln His Ile Asp Gln Leu Met Glu Met Glu Pro Gln Leu Glu Glu Leu
        115                 120                 125

His Pro Asp Glu Leu Glu Asn Ile Lys Ala His Arg Leu Ala Val Gln
    130                 135                 140

Glu Gln Leu Gln Arg Leu Gln Ala Pro Leu Asp Asp Arg Arg Lys Ala
145                 150                 155                 160

Leu Glu Arg Lys Lys Ala Ala Phe Gln Phe Gly Arg Asp Val Asp Asp
                165                 170                 175

Glu Lys Leu Trp Ile Ser Glu Arg Leu Val Leu Ala Lys Ala Gln Asn
            180                 185                 190

Leu Gly Glu Ser Leu Pro Asp Cys His Arg Leu Gln Lys Asn Leu Gln
        195                 200                 205

Leu Leu Ser Asn Glu Ile Asp Asn His Glu Pro Trp Ile Asn Gln Ile
    210                 215                 220
```

Cys Asn Asn Gly Gln Glu Leu Ile Asp Glu Gly His Ala Asn Gly Pro
225 230 235 240

Ala Phe Glu Lys Lys Ile Gln Glu Leu Arg Ser Ala Trp Gln Glu Leu
245 250 255

Lys Glu Ala Val Lys Asp Arg Lys Gly Asp Leu Gly Glu Ser Glu Lys
260 265 270

Ala His Gln Phe Leu Tyr Asp Cys Gly Glu Ala Glu Ala Trp Met Ser
275 280 285

Glu Gln Glu Leu Tyr Met Met Gln Asp Glu Arg Gly Lys Asp Glu Phe
290 295 300

Ser Thr Lys Asn Gln Ile Lys Lys His Glu Arg Leu Gln Ser Asp Ile
305 310 315 320

Asp Lys Phe Ala Asp Thr Ile Arg Ala Leu Ala Thr Lys Ala His Lys
325 330 335

Phe Val Glu Glu Lys Ser Pro Leu Thr Glu Gln Ile Gln Val Arg Gln
340 345 350

Ala Gln Ile Glu Lys Leu Tyr Ala Gly Leu Gln Asp Leu Ser Lys Glu
355 360 365

Arg Arg Lys Arg Leu Glu Glu Thr Leu Glu Leu Tyr Ala Leu His Arg
370 375 380

Glu Ile Asp Asp Leu Leu Gln Trp Ile Ala Asp Lys Glu Val Val Ala
385 390 395 400

Gly Ser Gln Glu Asn Gly Gln Asp Tyr Glu His Val Gln Met Leu Gln
405 410 415

Glu Arg Phe Gln Gln Phe Ala Arg Asp Thr Glu Asn Ile Gly Ser Glu
420 425 430

Arg Val Ala Asn Ala Asn Asp Gly Cys Asp Thr Leu Ile Gly His Gly
435 440 445

His Thr Asp Ala Pro Thr Ile Ala Leu Trp Lys Asp Ser Leu Asn Glu
450 455 460

Ala Trp Glu Asn Leu Leu Glu Leu Met Asp Thr Arg Ala Gln Ile Leu
465 470 475 480

Glu Ala Ser Arg Leu Leu His Lys Phe Tyr His Asp Cys Arg Asp Cys
485 490 495

Leu Ser Arg Ile Met Glu Lys Thr His Ala Met Pro Asp Asp Leu Gly
500 505 510

Arg Asp Ser Ser Ser Val Gly Ala Leu Ser Arg Lys His Gln Asn Tyr
515 520 525

Leu Lys Asp Ile Ala Ala Ile Gly Glu Gln Val Ala Gln Ile Glu Arg

530 535 540

Asp Ala Ala Glu Leu Arg Asp Gly Tyr Ala Gly Asp Lys Ala Leu Asp
545 550 555 560

Ile Gly Ser Arg Glu Ser Glu Val Val Lys Ala Trp Arg His Leu Arg
565 570 575

Gly Leu Cys Asp Ala Arg Thr Ser Arg Leu Met Asp Thr Ser Asp Leu
580 585 590

Phe Lys Phe Met Asn Met Val Arg Asp Leu Leu Leu Trp Met Asp Glu
595 600 605

Val Lys Arg Glu Met Asn Ser Gln Glu Arg Pro Lys Asp Val Ser Gly
610 615 620

Val Glu Leu Leu Met Asn Asn His Gln Ser Leu Lys Ala Glu Ile Asp
625 630 635 640

Ala Arg Glu Glu Asn Phe Asn Ala Cys Ile Ser Leu Gly Arg Asp Leu
645 650 655

Leu Asn Arg Lys His Tyr Ala Ser Ser Glu Ile Glu Lys Lys Leu Ile
660 665 670

Lys Leu Thr Thr Glu Arg Ala Glu Met Met Arg Arg Trp Glu Asp Arg
675 680 685

Trp Glu Tyr Leu Gln Leu Ile Leu Glu Val Tyr Gln Phe Ala Arg Asp
690 695 700

Ala Ala Val Ala Glu Ser Trp Leu Phe Ala Gln Glu Pro Tyr Leu Ile
705 710 715 720

Ser Lys Glu Tyr Gly Arg Asn Leu Glu Glu Thr
725 730

<210> 15
<211> 3561
<212> DNA
<213> Caenorhabditis elegans

<400> 15
atgagctcta gtagtagtga tgaaaatgaa acaacgatac acagaacagg atcaaatacc 60
ggtggaagtg gaatctactc acaacctcga gcaggatcat caaaaagaac atcgaatgtg 120
cgacatgatg ttagtgatgt agatgatgaa gaagagcatt atgcaagatt tcgcgaagat 180
acggcgatcg aggttgacga tgctataaca gttcttcttt catctctaca tttcgaacac 240
aaacgtgata ttgttccgac cgatgaagat gataataaac ttcgagaact tcacgaaaag 300
atttttgcat tgataacgag tgaatctgat gttaacagaa aaagacggct gaaaaaagct 360
cttcctgcgt caaactgtgt tagagaacaa gtatattatc ttcgaaggaa accatcaaca 420
ccaccagctt cttattatca tcgactaaat gcagctcttc acacaatcgt taaggaatca 480
tttggagaag aatatcgaaa agttgctaca gtattaggac ttgttgaagc gcttgcagaa 540
gttttaattc ttgaagttca cacatttgga atcaatgaga ccaatcctgg agagcatcga 600
aatattcgaa aacttattgc caatgctctt acgaatttaa catacggtca aattcactct 660
aaacgccggc tatgctcata tgatggattc attcgatgcg ttgttcgaat tgtgatcgaa 720
tctccaaata taactcaagt atacgctgga ctaatcagaa atctctcatg gaacgccgat 780

tcaggaatga gtgaagctct tcaaccaaca gttcatgcac tttcaattgc tgcagtacat 840
gcacatacac atagattcga tgttacagca acattatctg ctttgtggaa tcttgctggt 900
cactcagtag agaataaacg aactatttgt gatactccga attgtttaaa agtattagcc 960
agccttctat ctcctgatgc aagattcacg tcacttgttg attcagctac tggaattctg 1020
aagtatgtat cacaatatct cgcgaatact tccacgcatt tagagctaag aagcttgctc 1080
atcactcgaa tgctcacttt gctcaagagt gcctcgttca cttgtgttac aaatactcta 1140
ggagctattg caaacctaat tgtcaaagat ccccacatgc agcaaatgat tcggcaggat 1200
atggctgcag tacagcaact caacgtactc cgaaactcga atcgtgatga tattcggaca 1260
gcagtgaagt ctgtgctgaa tactcttaat caaccgtgta gtcatcgtta tggtgatatg 1320
tcacattctg ttggaggagg agccacaggg atgcaaatgc tttcggagcc acaacttcag 1380
atgcaaacat ctcatcatgc ttatcatgga actgcgtcac caagattact atcactgaga 1440
gcaactcgag catcacctgg aaaatatata caaccacagg ctcaacaaca gttgatacag 1500
actccacaag ttgatcaaag atctagtagt ctgccaagac attttgccgt tcaacgaaat 1560
ggatttgtga tggcacagtc ctacaatcaa caaatggatc aacatcagca acaacaaatg 1620
atctatcaat tgcagcaaca acaacaaatt atgacggagg atcaagcgca aatggaacat 1680
catcaacaaa tcatgtacct tcagcagcaa caacaacaat ttcatcaaat tcaacaacag 1740
caacaaatgc aaaaagctca ggaagctgat ccagttcctc ctacggatga cgatctagat 1800
attcccacat ctaccgtaat gggaacacga agtaacagtg aaagaagtct cggctcgatg 1860
aatcctggaa gtgttatgac aaactggaat tcttctctgg atactgcagc caacagttct 1920
cgtgctcttt ctcctgtatc atacaatgat attccagctt ctccaacaat gtgtgctcag 1980
gtgttcaatt tacctaaatc cactgaatcg gaacatcatc aattgacgtc tcaacagcag 2040
aacactacac attatagcag tggatcagct aacactatga ctcggagtga tggtgcaact 2100
actgtgccaa tggataacat tattacacca acttatgcta ttttgaatcc aatattggta 2160
cacgaacaga caccgaacgg gacagttcca cgaaagactt cagaagagtt ggatagtcct 2220
gatgatgttc tcccaggtcc atctcttgag gaggaagaag gtgattatgc tataatagga 2280
ggtgccgcac aaaagacaga tgacgagctt ttgactcgtt ctattcaatc tgaaatgcct 2340
acttcgtcgt ctactccgaa gatgaaagtc tcgcctcgcc ttaatggatt cttctctcca 2400
actcagaaaa caacatcgtc tcctgcatgg tctcatccag atacctcacc gattccaaag 2460
tcatcttctc atcgaactca acctaatcgt cgtcaagatg cttctgatgc agatcgtctt 2520
ttaatggaat caataatgtc agagatgcca aaaagtcgaa ttatatctcc acgacttgct 2580
ggtacacaac aatatcttga acccgagcca gagagaagaa gtcattcaaa aaacgaggaa 2640
gccgatcgcc gcgacgcatt caccgcatcg catgagccca gcgatcataa cggcattgat 2700
gtggcaagag gctcagactg gtctccacaa caacaattac atcgaatgga gtcattggaa 2760
agccaagcta gtagtgaaga tagttttgga ttaactgctg aagaaccaaa ttcatcaact 2820
tctggtgctg ctgcaaacac tatgagattt gacgatgaga tcgatgcatc tcttccaatg 2880
gattgtgtcg atgatgatga ttacgactat acttatgatc atttcgaaga ttatgaagat 2940
gaggaggatc ctgatgcgac tcagtttgat gatggagttg atgcacaact aactatcgat 3000
tgttcaatga tttctagtgg aagcggttca tcgcaaagaa atgaaactac gacaacgtca 3060
cgagattcaa aagctttagc tacttccaca ccaaaaggat ctgcatcttc actcccagga 3120
gttcgacaag ccacccgagt atcaacgaat ggaaagtcaa gactgcctgt tccaaaaacg 3180
aatggatcac ttgttgataa gaatcccaag ccaatcatcg cttcccggcg gccacgattg 3240
ccaccgaagc cgacactttt aaaagataag cactatccag aagaggattc gattgagaat 3300
cagacgagag atgatacgat atatgtgaat gctccagttg ttgaggcaga acaagaacga 3360
atttatatga atgcattgaa gcaacagaag aacattgagc aatcaccgtc aattggaaac 3420
ggatcaccta tagctaaatc tgcaattgta acgccctaca attatcagaa accgcctttc 3480
actggtcgaa ataatggtga aatgagcaac gagaagagtg tgacaccgaa tcctaaacaa 3540
atgcttgtaa caatagtcta a 3561

<210> 16
<211> 1186
<212> PRT
<213> Caenorhabditis elegans

<400> 16
Met Ser Ser Ser Ser Ser Asp Glu Asn Glu Thr Thr Ile His Arg Thr
1 5 10 15

Gly Ser Asn Thr Gly Gly Ser Gly Ile Tyr Ser Gln Pro Arg Ala Gly

20 25 30

Ser Ser Lys Arg Thr Ser Asn Val Arg His Asp Val Ser Asp Val Asp
35 40 45

Asp Glu Glu Glu His Tyr Ala Arg Phe Arg Glu Asp Thr Ala Ile Glu
50 55 60

Val Asp Asp Ala Ile Thr Val Leu Leu Ser Ser Leu His Phe Glu His
65 70 75 80

Lys Arg Asp Ile Val Pro Thr Asp Glu Asp Asp Asn Lys Leu Arg Glu
85 90 95

Leu His Glu Lys Ile Phe Ala Leu Ile Thr Ser Glu Ser Asp Val Asn
100 105 110

Arg Lys Arg Arg Leu Lys Lys Ala Leu Pro Ala Ser Asn Cys Val Arg
115 120 125

Glu Gln Val Tyr Tyr Leu Arg Arg Lys Pro Ser Thr Pro Pro Ala Ser
130 135 140

Tyr Tyr His Arg Leu Asn Ala Ala Leu His Thr Ile Val Lys Glu Ser
145 150 155 160

Phe Gly Glu Glu Tyr Arg Lys Val Ala Thr Val Leu Gly Leu Val Glu
165 170 175

Ala Leu Ala Glu Val Leu Ile Leu Glu Val His Thr Phe Gly Ile Asn
180 185 190

Glu Thr Asn Pro Gly Glu His Arg Asn Ile Arg Lys Leu Ile Ala Asn
195 200 205

Ala Leu Thr Asn Leu Thr Tyr Gly Gln Ile His Ser Lys Arg Arg Leu
210 215 220

Cys Ser Tyr Asp Gly Phe Ile Arg Cys Val Val Arg Ile Val Ile Glu
225 230 235 240

Ser Pro Asn Ile Thr Gln Val Tyr Ala Gly Leu Ile Arg Asn Leu Ser
245 250 255

Trp Asn Ala Asp Ser Gly Met Ser Glu Ala Leu Gln Pro Thr Val His
260 265 270

Ala Leu Ser Ile Ala Ala Val His Ala His Thr His Arg Phe Asp Val
275 280 285

Thr Ala Thr Leu Ser Ala Leu Trp Asn Leu Ala Gly His Ser Val Glu
290 295 300

Asn Lys Arg Thr Ile Cys Asp Thr Pro Asn Cys Leu Lys Val Leu Ala
305 310 315 320

Ser Leu Leu Ser Pro Asp Ala Arg Phe Thr Ser Leu Val Asp Ser Ala
325 330 335

Thr Gly Ile Leu Lys Tyr Val Ser Gln Tyr Leu Ala Asn Thr Ser Thr
340 345 350

His Leu Glu Leu Arg Ser Leu Leu Ile Thr Arg Met Leu Thr Leu Leu
355 360 365

Lys Ser Ala Ser Phe Thr Cys Val Thr Asn Thr Leu Gly Ala Ile Ala
370 375 380

Asn Leu Ile Val Lys Asp Pro His Met Gln Gln Met Ile Arg Gln Asp
385 390 395 400

Met Ala Ala Val Gln Gln Leu Asn Val Leu Arg Asn Ser Asn Arg Asp
405 410 415

Asp Ile Arg Thr Ala Val Lys Ser Val Leu Asn Thr Leu Asn Gln Pro
420 425 430

Cys Ser His Arg Tyr Gly Asp Met Ser His Ser Val Gly Gly Gly Ala
435 440 445

Thr Gly Met Gln Met Leu Ser Glu Pro Gln Leu Gln Met Gln Thr Ser
450 455 460

His His Ala Tyr His Gly Thr Ala Ser Pro Arg Leu Leu Ser Leu Arg
465 470 475 480

Ala Thr Arg Ala Ser Pro Gly Lys Tyr Ile Gln Pro Gln Ala Gln Gln
485 490 495

Gln Leu Ile Gln Thr Pro Gln Val Asp Gln Arg Ser Ser Ser Leu Pro
500 505 510

Arg His Phe Ala Val Gln Arg Asn Gly Phe Val Met Ala Gln Ser Tyr
515 520 525

Asn Gln Gln Met Asp Gln His Gln Gln Gln Gln Met Ile Tyr Gln Leu
530 535 540

Gln Gln Gln Gln Gln Ile Met Thr Glu Asp Gln Ala Gln Met Glu His
545 550 555 560

His Gln Gln Ile Met Tyr Leu Gln Gln Gln Gln Gln Gln Phe His Gln
565 570 575

Ile Gln Gln Gln Gln Gln Met Gln Lys Ala Gln Glu Ala Asp Pro Val
580 585 590

Pro Pro Thr Asp Asp Asp Leu Asp Ile Pro Thr Ser Thr Val Met Gly
595 600 605

Thr Arg Ser Asn Ser Glu Arg Ser Leu Gly Ser Met Asn Pro Gly Ser
610 615 620

Val Met Thr Asn Trp Asn Ser Ser Leu Asp Thr Ala Ala Asn Ser Ser
625 630 635 640

Arg Ala Leu Ser Pro Val Ser Tyr Asn Asp Ile Pro Ala Ser Pro Thr
645 650 655

Met Cys Ala Gln Val Phe Asn Leu Pro Lys Ser Thr Glu Ser Glu His
660 665 670

His Gln Leu Thr Ser Gln Gln Gln Asn Thr Thr His Tyr Ser Ser Gly
675 680 685

Ser Ala Asn Thr Met Thr Arg Ser Asp Gly Ala Thr Thr Val Pro Met
690 695 700

Asp Asn Ile Ile Thr Pro Thr Tyr Ala Ile Leu Asn Pro Ile Leu Val
705 710 715 720

His Glu Gln Thr Pro Asn Gly Thr Val Pro Arg Lys Thr Ser Glu Glu
725 730 735

Leu Asp Ser Pro Asp Asp Val Leu Pro Gly Pro Ser Leu Glu Glu Glu
740 745 750

Glu Gly Asp Tyr Ala Ile Ile Gly Gly Ala Ala Gln Lys Thr Asp Asp
755 760 765

Glu Leu Leu Thr Arg Ser Ile Gln Ser Glu Met Pro Thr Ser Ser Ser
770 775 780

Thr Pro Lys Met Lys Val Ser Pro Arg Leu Asn Gly Phe Phe Ser Pro
785 790 795 800

Thr Gln Lys Thr Thr Ser Ser Pro Ala Trp Ser His Pro Asp Thr Ser
805 810 815

Pro Ile Pro Lys Ser Ser Ser His Arg Thr Gln Pro Asn Arg Arg Gln
820 825 830

Asp Ala Ser Asp Ala Asp Arg Leu Leu Met Glu Ser Ile Met Ser Glu
835 840 845

Met Pro Lys Ser Arg Ile Ile Ser Pro Arg Leu Ala Gly Thr Gln Gln
850 855 860

Tyr Leu Glu Pro Glu Pro Glu Arg Arg Ser His Ser Lys Asn Glu Glu
865 870 875 880

Ala Asp Arg Arg Asp Ala Phe Thr Ala Ser His Glu Pro Ser Asp His
885 890 895

Asn Gly Ile Asp Val Ala Arg Gly Ser Asp Trp Ser Pro Gln Gln Gln
900 905 910

Leu His Arg Met Glu Ser Leu Glu Ser Gln Ala Ser Ser Glu Asp Ser
915 920 925

Phe Gly Leu Thr Ala Glu Glu Pro Asn Ser Ser Thr Ser Gly Ala Ala
930 935 940

Ala Asn Thr Met Arg Phe Asp Asp Glu Ile Asp Ala Ser Leu Pro Met

945 950 955 960

Asp Cys Val Asp Asp Asp Asp Tyr Asp Tyr Thr Tyr Asp His Phe Glu
965 970 975

Asp Tyr Glu Asp Glu Glu Asp Pro Asp Ala Thr Gln Phe Asp Asp Gly
980 985 990

Val Asp Ala Gln Leu Thr Ile Asp Cys Ser Met Ile Ser Ser Gly Ser
995 1000 1005

Gly Ser Ser Gln Arg Asn Glu Thr Thr Thr Thr Ser Arg Asp Ser Lys
1010 1015 1020

Ala Leu Ala Thr Ser Thr Pro Lys Gly Ser Ala Ser Ser Leu Pro Gly
1025 1030 1035 1040

Val Arg Gln Ala Thr Arg Val Ser Thr Asn Gly Lys Ser Arg Leu Pro
1045 1050 1055

Val Pro Lys Thr Asn Gly Ser Leu Val Asp Lys Asn Pro Lys Pro Ile
1060 1065 1070

Ile Ala Ser Arg Arg Pro Arg Leu Pro Pro Lys Pro Thr Leu Leu Lys
1075 1080 1085

Asp Lys His Tyr Pro Glu Glu Asp Ser Ile Glu Asn Gln Thr Arg Asp
1090 1095 1100

Asp Thr Ile Tyr Val Asn Ala Pro Val Val Glu Ala Glu Gln Glu Arg
1105 1110 1115 1120

Ile Tyr Met Asn Ala Leu Lys Gln Gln Lys Asn Ile Glu Gln Ser Pro
1125 1130 1135

Ser Ile Gly Asn Gly Ser Pro Ile Ala Lys Ser Ala Ile Val Thr Pro
1140 1145 1150

Tyr Asn Tyr Gln Lys Pro Pro Phe Thr Gly Arg Asn Asn Gly Glu Met
1155 1160 1165

Ser Asn Glu Lys Ser Val Thr Pro Asn Pro Lys Gln Met Leu Val Thr
1170 1175 1180

Ile Val
1185

<210> 17
<211> 1742
<212> DNA
<213> Caenorhabditis elegans

<400> 17
atgagctcta gtagtagtga tgaaaatgaa acaacgatac acagaacagg atcaaatacc 60
ggtggaagtg gaatctactc acaacctcga gcaggatcat caaaaagaac atcgaatgtg 120
cgacatgatg ttagtgatgt agatgatgaa gaagagcatt atgcaagatt tcgcgaagat 180
acggcgatcg aggttgacga tgctataaca gttcttcttt catctctaca tttcgaacac 240

aaacgtgata ttgttccgac cgatgaagat gataataaac ttcgagaact tcacgaaaag 300
attttgcat tgataacgag tgaatctgat gttaacagaa aaagacggct gaaaaaagct 360
cttcctgcgt caaactgtgt tagagaacaa gtatattatc ttcgaaggaa accatcaaca 420
ccaccagctt cttattatca tcgactaaat gcagctcttc acacaatcgt taaggaatca 480
tttggagaag aatatcgaaa agttgctaca gtattaggac ttgttgaagc gcttgcagaa 540
gttttaattc ttgaagttca cacatttgga atcaatgaga ccaatcctgg agagcatcga 600
aatattcgaa aacttattgc caatgctctt acgaatttaa catacggtca aattcactct 660
aaacgccggc tatgctcata tgatggattc attcgatgcg ttgttcgaat tgtgatcgaa 720
tctccaaata taactcaagt atacgctgga ctaatcagaa atctctcatg gaacgccgat 780
tcaggaatga gtgaagctct tcaaccaaca gttcatgcac tttcaattgc tgcagtacat 840
gcacatacac atagattcga tgttacagca acattatctg ctttgtggaa tcttgctggt 900
cactcagtag agaataaacg aactatttgt gatactccga attgtttaaa agtattagcc 960
agccttctat ctcctgatgc aagattcacg tcacttgttg attcagctac tggaattctg 1020
aagtatgtat cacaatatct cgcgaatact tccacgcatt tagagctaag aagcttgctc 1080
atcactcgaa tgctcacttt gctcaagagt gcctcgttca cttgtgttac aaatactcta 1140
ggagctattg caaacctaat tgtcaaagat ccccacatgc agcaaatgat tcggcaggat 1200
atggctgcag tacagcaact caacgtactc cgaaactcga atcgtgatga tattcggaca 1260
gcagtgaagt ctgtgctgaa tactcttaat caaccgtgta gtcatcgtta tggtgatatg 1320
tcacattctg ttggaggagg agccacaggg atgcaaatgc tttcggagcc acaacttcag 1380
atgcaaacat ctcatcatgc ttatcatgga actgcgtcac caagattact atcactgaga 1440
gcaactcgag catcacctgg aaaatatata caaccacagg ctcaacaaca gttgatacag 1500
actccacaag ttgatcaaag atctagtagt ctgccaagac attttgccgt tcaacgaaat 1560
ggatttgtga tggcacagtc ctacaatcaa caaatggatc aacatcagca acaacaaatg 1620
atctatcaat tgcagcaaca acaacaaatt atgacggagg atcaagcgca aatggaacat 1680
catcaacaaa tcatgtacct tcagcagcaa caacaacaat ttcatcaaat tcaacaacag 1740
ca 1742

<210> 18
<211> 509
<212> PRT
<213> Caenorhabditis elegans

<400> 18
Met Ser Ser Ser Ser Ser Asp Glu Asn Glu Thr Thr Ile His Arg Thr
1 5 10 15

Gly Ser Asn Thr Gly Gly Ser Gly Ile Tyr Ser Gln Pro Arg Ala Gly
20 25 30

Ser Ser Lys Arg Thr Ser Asn Val Arg His Asp Val Ser Asp Val Asp
35 40 45

Asp Glu Glu Glu His Tyr Ala Arg Phe Arg Glu Asp Thr Ala Ile Glu
50 55 60

Val Asp Asp Ala Ile Thr Val Leu Leu Ser Ser Leu His Phe Glu His
65 70 75 80

Lys Arg Asp Ile Val Pro Thr Asp Glu Asp Asp Asn Lys Leu Arg Glu
85 90 95

Leu His Glu Lys Ile Phe Ala Leu Ile Thr Ser Glu Ser Asp Val Asn
100 105 110

Arg Lys Arg Arg Leu Lys Lys Ala Leu Pro Ala Ser Asn Cys Val Arg
115 120 125

Glu Gln Val Tyr Tyr Leu Arg Arg Lys Pro Ser Thr Pro Pro Ala Ser

130 135 140

Tyr Tyr His Arg Leu Asn Ala Ala Leu His Thr Ile Val Lys Glu Ser
145 150 155 160

Phe Gly Glu Glu Tyr Arg Lys Val Ala Thr Val Leu Gly Leu Val Glu
165 170 175

Ala Leu Ala Glu Val Leu Ile Leu Glu Val His Thr Phe Gly Ile Asn
180 185 190

Glu Thr Asn Pro Gly Glu His Arg Asn Ile Arg Lys Leu Ile Ala Asn
195 200 205

Ala Leu Thr Asn Leu Thr Tyr Gly Gln Ile His Ser Lys Arg Arg Leu
210 215 220

Cys Ser Tyr Asp Gly Phe Ile Arg Cys Val Val Arg Ile Val Ile Glu
225 230 235 240

Ser Pro Asn Ile Thr Gln Val Tyr Ala Gly Leu Ile Arg Asn Leu Ser
245 250 255

Trp Asn Ala Asp Ser Gly Met Ser Glu Ala Leu Gln Pro Thr Val His
260 265 270

Ala Leu Ser Ile Ala Ala Val His Ala His Thr His Arg Phe Asp Val
275 280 285

Thr Ala Thr Leu Ser Ala Leu Trp Asn Leu Ala Gly His Ser Val Glu
290 295 300

Asn Lys Arg Thr Ile Cys Asp Thr Pro Asn Cys Leu Lys Val Leu Ala
305 310 315 320

Ser Leu Leu Ser Pro Asp Ala Arg Phe Thr Ser Leu Val Asp Ser Ala
325 330 335

Thr Gly Ile Leu Lys Tyr Val Ser Gln Tyr Leu Ala Asn Thr Ser Thr
340 345 350

His Leu Glu Leu Arg Ser Leu Leu Ile Thr Arg Met Leu Thr Leu Leu
355 360 365

Lys Ser Ala Ser Phe Thr Cys Val Thr Asn Thr Leu Gly Ala Ile Ala
370 375 380

Asn Leu Ile Val Lys Asp Pro His Met Gln Gln Met Ile Arg Gln Asp
385 390 395 400

Met Ala Ala Val Gln Gln Leu Asn Val Leu Arg Asn Ser Asn Arg Asp
405 410 415

Asp Ile Arg Thr Ala Val Lys Ser Val Leu Asn Thr Leu Asn Gln Pro
420 425 430

Cys Ser His Arg Tyr Gly Asp Met Ser His Ser Val Gly Gly Gly Ala
435 440 445

```
Thr Gly Met Gln Met Leu Ser Glu Pro Gln Leu Gln Met Gln Thr Ser
    450                 455                     460

His His Ala Tyr His Gly Thr Ala Ser Pro Arg Leu Leu Ser Leu Arg
465                 470                     475                 480

Ala Thr Arg Ala Ser Pro Gly Lys Tyr Ile Gln Pro Gln Ala Gln Gln
                485                     490                 495

Glu Leu Ile Gln Thr Pro Gln Val Asp Gln Arg Ser Ser
            500                     505


<210> 19
<211> 1998
<212> DNA
<213> Caenorhabditis elegans

<400> 19
atgatggaac tgtgcgaact gcacatgatg cctgaagagg atcaatggcc cgatgacatt 60
aatgctcgaa tcgctgctct caacttttgt ggagaagatt cgatgctctt tgaatcagtt 120
gatcctttca tatcaactga ttctttagac tctcaacagt ttcgagagag atgtcaaatg 180
aaaaaagaag attttcaatt agcattcgca gactctggac attggcaatc aggtataaat 240
gataatctga caacttgggg tcggattcgg acatctgaac cattggatga acgtactgcc 300
agtgctccag atgtttggaa tgtaaaaaga agtgatagtg ctcgatcacc aaatcgtcca 360
aattcactga ttgccaactt tgtttccgga gatgctactc gattcgtcga tgtaaatgat 420
aatgaaattc gcgaggcaaa tgaagaaata attcgtaaag atcgttggag aagagattca 480
acgagacgat gctcatctgg aggacaaaat cagaaacgaa catttgcgga tattctggaa 540
aagaatgtaa cagcaccaac ttcaatggca atcacatcgt ctgataatga aaagcctccg 600
aaattagact tcctggcaat gcatcatgag atgccatctc tctgcgagtc attcaccgct 660
tcgttccgtg atgctattat taaaatgcag aaatgtgagc cactaccgtc gattacttca 720
accaatgatt ttccgctctt ctttcaagaa gactctcccg actcaggact cggttgtagt 780
ggtccgagcc acattgaaga ttggcaatct ctatctgttc ttctgcccaa acacgtcgca 840
gaagcatgtt cattcttcaa gagcaacacc caattattaa catcaagtac atcaaaaaca 900
gcaccccaaa catcaacaaa catcgtatca aattgcattg atcgccgtat ctctggtatt 960
tcagcatcag caaatgaagc atgtcgaacg tgttatcgag tacgaaggag gatacaccca 1020
ccagtttggc ctcaaactgc ccaatcaaag acggttcttt gtgactgcgc atcaacaccc 1080
acggatacga atttctcatt tgcaccaaca acttcaacga ctcgccatca actgcgcgca 1140
aaagagctct caattgtcgg attgccgatt tatgcagcaa agcggacact cgttgagaat 1200
gttgttgaag gcgttgctgc aatttcaaga ggggatggaa gtgatttgct tgtcatcgcc 1260
atgagatgct taattgaaga tggtcttcaa gaaaatgtat cggcatggac aatgattcag 1320
acggttacct caaaaggccc agctacaaaa gatgttcatt caattgtaaa acaattggaa 1380
gaatgttcaa agactgacaa cgtcaaagtt gaaatcttct ttgaggagct tatcagagaa 1440
aactcattgg actgctggct gtgctatatt gtgttgaaag agaaggtgct gaaaacattg 1500
tattctgaaa atgcattcct tctttctgcg agttctgaat acagaactct tctttggaga 1560
atggtcgaca gtttgtctct tttgccagtg atcgaagctc gcagtgattc agttcaccaa 1620
cagttcaaat caatgcaaca atggggcgga gcatccagaa tagcttctga ctcaagagtt 1680
ccaaagagtt catcgttccc agcaagatta tcaactgctc catcaagaag atcaaggata 1740
ccattgtcaa caagtcgcat cagtatcagc tcaacaactt caacacctcg atcagctcga 1800
tcaccttcaa cgacatctcg tattcgagtt gcatctataa tgggtgactt cacacttgcc 1860
aatttctccc tctcagacgg agaaaaagtt agtgttctat caactcgtgg aggacttgct 1920
cgatgtgtcc gtttaacgac ttctcattca aaaatcaaca atggagttat tccgattgaa 1980
catcttcttt tccagtga                                               1998

<210> 20
<211> 665
<212> PRT
```

<213> Caenorhabditis elegans

<400> 20
Met Val Glu Leu Cys Glu Leu His Met Met Pro Glu Glu Asp Gln Trp
1 5 10 15

Pro Asp Asp Ile Asp Ala Arg Ile Ala Ala Leu Asn Phe Cys Gly Glu
20 25 30

Asp Ser Met Leu Phe Glu Ser Val Asp Pro Ser Val Ser Thr Asp Ser
35 40 45

Leu Asp Ser Gln Gln Phe Arg Glu Arg Cys Gln Met Lys Lys Glu Asp
50 55 60

Phe Gln Leu Ala Phe Ala Asp Ser Gly His Trp Gln Ser Gly Ile Asn
65 70 75 80

Asp Asn Leu Thr Thr Trp Gly Arg Ile Arg Thr Ser Glu Pro Leu Asp
85 90 95

Glu Arg Thr Ala Ser Ala Pro Asp Val Trp Asn Val Lys Arg Ser Asp
100 105 110

Ser Ala Arg Ser Pro Asn Arg Pro Asn Ser Leu Ile Ala Asn Phe Val
115 120 125

Ser Gly Asp Ala Thr Arg Phe Val Asp Val Asn Asp Asn Glu Ile Arg
130 135 140

Glu Ala Asn Glu Glu Ile Ile Arg Lys Asp Arg Trp Arg Arg Asp Ser
145 150 155 160

Ala Arg Arg Cys Ser Ser Gly Gly Gln Asn Gln Lys Arg Thr Phe Ala
165 170 175

Asp Ile Leu Glu Lys Asn Val Thr Ala Pro Thr Ser Met Ala Ile Thr
180 185 190

Ser Ser Asp Asn Glu Lys Pro Pro Lys Leu Asp Phe Leu Ala Met His
195 200 205

His Glu Met Pro Ser Leu Cys Glu Ser Phe Thr Ala Ser Phe Arg Asp
210 215 220

Ala Ile Ile Lys Met Gln Lys Cys Glu Pro Leu Pro Ser Ile Thr Ser
225 230 235 240

Thr Asn Asp Phe Pro Leu Phe Phe Gln Glu Asp Ser Pro Asp Ser Gly
245 250 255

Leu Gly Cys Ser Gly Pro Ser His Ile Glu Asp Trp Gln Ser Leu Ser
260 265 270

Val Leu Leu Pro Lys His Val Ala Glu Ala Cys Ser Phe Phe Lys Ser
275 280 285

Asn Thr Gln Leu Leu Thr Ser Ser Thr Ser Lys Thr Ala Pro Gln Thr

290 295 300

Ser Thr Asn Ile Val Ser Asn Cys Ile Asp Arg Arg Ile Ser Gly Ile
305 310 315 320

Ser Ala Ser Ala Asn Glu Ala Cys Arg Thr Cys Tyr Arg Val Arg Arg
325 330 335

Arg Ile His Pro Pro Val Trp Ala Gln Thr Ala Gln Ser Lys Thr Val
340 345 350

Leu Cys Asp Cys Ala Ser Thr Pro Thr Asp Thr Asn Phe Ser Phe Ala
355 360 365

Pro Thr Thr Ser Thr Thr Arg His Gln Leu Arg Ala Lys Glu Leu Ser
370 375 380

Ile Val Gly Leu Pro Ile Tyr Ala Ala Lys Arg Thr Leu Val Glu Asn
385 390 395 400

Val Val Glu Gly Val Ala Ala Ile Ser Arg Gly Asp Gly Ser Asp Leu
405 410 415

Leu Val Ile Ala Met Arg Cys Leu Ile Glu Asp Gly Leu Gln Glu Asn
420 425 430

Val Ser Ala Trp Thr Met Ile Gln Thr Val Thr Ser Lys Gly Pro Ala
435 440 445

Thr Lys Asp Val His Ser Ile Val Lys Gln Leu Glu Glu Cys Ser Lys
450 455 460

Thr Asp Asn Val Lys Val Glu Ile Phe Phe Glu Glu Leu Ile Arg Glu
465 470 475 480

Asn Ser Leu Asp Cys Trp Leu Cys Tyr Ile Val Leu Lys Glu Lys Val
485 490 495

Leu Lys Thr Leu Tyr Ser Glu Asn Ala Phe Leu Leu Ser Ala Ser Ser
500 505 510

Glu Tyr Arg Thr Leu Leu Trp Arg Met Val Asp Ser Leu Ser Leu Leu
515 520 525

Pro Val Ile Glu Ala Arg Ser Asp Ser Val His Gln Gln Phe Lys Ser
530 535 540

Met Gln Gln Trp Gly Gly Ala Ser Arg Ile Ala Ser Asp Ser Arg Val
545 550 555 560

Pro Lys Ser Ser Ser Phe Pro Ala Arg Leu Ser Thr Ala Pro Ser Arg
565 570 575

Arg Ser Arg Ile Pro Leu Ser Thr Ser Arg Ile Ser Ile Ser Ser Thr
580 585 590

Thr Ser Thr Pro Arg Ser Ala Arg Ser Pro Ser Thr Thr Ser Arg Ile
595 600 605

Arg Val Ala Ser Ile Met Gly Asp Phe Thr Leu Ala Asn Phe Ser Leu
610 615 620

Ser Asp Gly Glu Lys Val Ser Val Leu Ser Thr Arg Gly Gly Leu Ala
625 630 635 640

Arg Cys Val Arg Leu Thr Thr Ser His Ser Lys Ile Asn Asn Gly Val
645 650 655

[illegible] [illegible] Ile Glu His Leu Leu Phe Gln
660 665

<210> 21
<211> 1839
<212> DNA
<213> Caenorhabditis elegans

<400> 21
[illegible] gatgtcaaat gaaaaaagaa gattttcaat tagcattcgc agactctgga 60
[illegible] caggtataaa tgataatctg acaacttggg gtcggattcg gacatctgaa 120
ccattggatg aacgtactgc cagtgctcca gatgtttgga atgtaaaaag aagtgatagt 180
gctcgatcac caaatcgtcc aaattcactg attgccaact ttgtttccgg agatgctact 240
cgatttgtcg atgtaaatga taatgaaatt cgcgaggcaa atgaagaaat aattcgtaaa 300
gatcgttgga gaagagattc agcgagacga tgctcatctg gaggacaaaa tcagaaacga 360
acatttgcgg atattctgga aaagaatgta acagcaccaa cttcaatggc aatcacatcg 420
tctgataatg aaaagcctcc gaaattagac ttcctggcaa tgcatcatga gatgccatct 480
ctctgcgagt cattcaccgc ttcgttccgt gatgctatta ttaaaatgca gaaatgtgag 540
ccactaccgt cgattacttc aaccaatgat tttccgctct tctttcaaga agactctccc 600
gactcaggac tcggttgtag tggtccgagc cacattgaag attggcaatc tctatctgtt 660
ctttctgccca aacacgtcgc agaagcatgt tcattcttca agagcaacac ccaattatta 720
acatcaagta catcaaaaac agcaccccaa acatcaacaa acatcgtatc aaattgcatt 780
gatcgccgta tctctggtat ttcagcatca gcaaatgaag catgtcgaac gtgttatcga 840
gtacgaagga ggatacaccc accagtttgg gctcaaactg cccaatcaaa gacggttctt 900
tgtgacttgcg catcaacacc cacggatacg aatttctcat ttgcaccaac aacttcaacg 960
actcgccatc aactgcgcgc aaaagagctc tcaattgtcg gattgccgat ttatgcagca 1020
aagcggacac tcgttgagaa tgttgttgaa ggcgttgctg caatttcaag aggggatgga 1080
agtgatttgc ttgtcatcgc catgagatgc ttaattgaag atggtcttca agaaaatgta 1140
tcggcatgga caatgattca gacggttacc tcaaaaggcc cagctacaaa agatgttcat 1200
tcaattgtaa aacaattgga agaatgttca aagactgaca acgtcaaagt tgaaatcttc 1260
tttgaggagc ttatcagaga aaactcattg gactgctggc tgtgctatat tgtgttgaaa 1320
gagaaggtgc tgaaaacatt gtattctgaa aatgcattcc ttctttctgc gagttctgaa 1380
tacagaactc ttctttggag aatggtcgac agtttgtctc ttttgccagt gatcgaagct 1440
cgcagtgatt cagttcacca acagttcaaa tcaatgcaac aatggggcgg agcatccaga 1500
atagcttctg actcaagagt tccaaagagt tcatcgttcc cagcaagatt atcaactgct 1560
ccatcaagaa gatcaaggat accattgtca acaagtcgca tcagtatcag ctcaacaact 1620
tcaacacctc gatcagctcg atcaccttca acgacatctc gtattcgagt tgcatctata 1680
atgggtgact tcacacttgc caatttctcc ctctcagacg gagaaaaagt tagtgttcta 1740
tcaactcgtg gaggacttgc tcgatgtgtc cgtttaacga cttctcattc aaaaatcaac 1800
aatggagtta ttccgattga acatcttctt ttccagtga 1839

<210> 22
<211> 612
<212> PRT
<213> Caenorhabditis elegans

<400> 22

Phe Arg Glu Arg Cys Gln Met Lys Lys Glu Asp Phe Gln Leu Ala Phe
1 5 10 15

Ala Asp Ser Gly His Trp Gln Ser Gly Ile Asn Asp Asn Leu Thr Thr
20 25 30

Trp Gly Arg Ile Arg Thr Ser Glu Pro Leu Asp Glu Arg Thr Ala Ser
35 40 45

Ala Pro Asp Val Trp Asn Val Lys Arg Ser Asp Ser Ala Arg Ser Pro
50 55 60

Asn Arg Pro Asn Ser Leu Ile Ala Asn Phe Val Ser Gly Asp Ala Thr
65 70 75 80

Arg Phe Val Asp Val Asn Asp Asn Glu Ile Arg Glu Ala Asn Glu Glu
85 90 95

Ile Ile Arg Lys Asp Arg Trp Arg Arg Asp Ser Ala Arg Arg Cys Ser
100 105 110

Ser Gly Gly Gln Asn Gln Lys Arg Thr Phe Ala Asp Ile Leu Glu Lys
115 120 125

Asn Val Thr Ala Pro Thr Ser Met Ala Ile Thr Ser Ser Asp Asn Glu
130 135 140

Lys Pro Pro Lys Leu Asp Phe Leu Ala Met His His Glu Met Pro Ser
145 150 155 160

Leu Cys Glu Ser Phe Thr Ala Ser Phe Arg Asp Ala Ile Ile Lys Met
165 170 175

Gln Lys Cys Glu Pro Leu Pro Ser Ile Thr Ser Thr Asn Asp Phe Pro
180 185 190

Leu Phe Phe Gln Glu Asp Ser Pro Asp Ser Gly Leu Gly Cys Ser Gly
195 200 205

Pro Ser His Ile Glu Asp Trp Gln Ser Leu Ser Val Leu Leu Pro Lys
210 215 220

His Val Ala Glu Ala Cys Ser Phe Phe Lys Ser Asn Thr Gln Leu Leu
225 230 235 240

Thr Ser Ser Thr Ser Lys Thr Ala Pro Gln Thr Ser Thr Asn Ile Val
245 250 255

Ser Asn Cys Ile Asp Arg Arg Ile Ser Gly Ile Ser Ala Ser Ala Asn
260 265 270

Glu Ala Cys Arg Thr Cys Tyr Arg Val Arg Arg Arg Ile His Pro Pro
275 280 285

Val Trp Ala Gln Thr Ala Gln Ser Lys Thr Val Leu Cys Asp Cys Ala
290 295 300

Ser Thr Pro Thr Asp Thr Asn Phe Ser Phe Ala Pro Thr Thr Ser Thr

305 310 315 320

Thr Arg His Gln Leu Arg Ala Lys Glu Leu Ser Ile Val Gly Leu Pro
325 330 335

Ile Tyr Ala Ala Lys Arg Thr Leu Val Glu Asn Val Val Glu Gly Val
340 345 350

Ala Ala Ile Ser Arg Gly Asp Gly Ser Asp Leu Leu Val Ile Ala Met
355 360 365

Arg Cys Leu Ile Glu Asp Gly Leu Gln Glu Asn Val Ser Ala Trp Thr
370 375 380

Met Ile Gln Thr Val Thr Ser Lys Gly Pro Ala Thr Lys Asp Val His
385 390 395 400

Ser Ile Val Lys Gln Leu Glu Glu Cys Ser Lys Thr Asp Asn Val Lys
405 410 415

Val Glu Ile Phe Phe Glu Glu Leu Ile Arg Glu Asn Ser Leu Asp Cys
420 425 430

Trp Leu Cys Tyr Ile Val Leu Lys Glu Lys Val Leu Lys Thr Leu Tyr
435 440 445

Ser Glu Asn Ala Phe Leu Leu Ser Ala Ser Ser Glu Tyr Arg Thr Leu
450 455 460

Leu Trp Arg Met Val Asp Ser Leu Ser Leu Leu Pro Val Ile Glu Ala
465 470 475 480

Arg Ser Asp Ser Val His Gln Gln Phe Lys Ser Met Gln Gln Trp Gly
485 490 495

Gly Ala Ser Arg Ile Ala Ser Asp Ser Arg Val Pro Lys Ser Ser Ser
500 505 510

Phe Pro Ala Arg Leu Ser Thr Ala Pro Ser Arg Arg Ser Arg Ile Pro
515 520 525

Leu Ser Thr Ser Arg Ile Ser Ile Ser Ser Thr Thr Ser Thr Pro Arg
530 535 540

Ser Ala Arg Ser Pro Ser Thr Thr Ser Arg Ile Arg Val Ala Ser Ile
545 550 555 560

Met Gly Asp Phe Thr Leu Ala Asn Phe Ser Leu Ser Asp Gly Glu Lys
565 570 575

Val Ser Val Leu Ser Thr Arg Gly Gly Leu Ala Arg Cys Val Arg Leu
580 585 590

Thr Thr Ser His Ser Lys Ile Asn Asn Gly Val Ile Pro Ile Glu His
595 600 605

Leu Leu Phe Gln
610

<210> 23
<211> 3435
<212> DNA
<213> Caenorhabditis elegans

<400> 23
atgcaggcga tagatgaatc caagcgaaat caaaaagtgc cgcctgcaaa acgaaagcgg 60
atatatttgt ctgacgagga agaggaagat tttgctgaag cggcgcacgt tgagaacaca 120
gtccccgagc gtgcaactcg acgtagcacc agaagacgtt catcaatgca cgaagaactt 180
ggagtttcag aacaggaaga gtcgccggtt cgcaggactc gcaaagccgc caagcgcctt 240
ggaagtgagc aaccagagga aaatctagca gcagatgatc cgttgccaat ggaaggaggt 300
ggcgagattg tgctaccaat tgctgaaatt gatggaatgg ctgaacaaga aaatgaggac 360
ttgattgaga aaattggaag ggaagaagag gaggaaggcg ctgaagaaga tgagcaaagc 420
ggtgaaaaag atccagaaga agaagaggat gatagcagca atgcagaaag ttcagaggaa 480
tcaactgctc cacgtcaata cagtcttcgc cgccgtcaac cagttgtcca gttcaatgca 540
tcagaagctc gtgaaaatcg acgggctcgt cttgaacatc atcgggttgc aaatcaaaac 600
agacatcacc gcaatcgaaa tggatctcgt cgaagacgaa gcgattcgga ttctgatagt 660
gatgatatgg tgcttccgag acctgataaa cgtcaatcga ggcctcatat gcataaccgt 720
ggagaacgtg aacgtggtcg atttatgccg atcaatatga ccgaaaaaga attgcagtct 780
gcccaacaca ttctgatgga tcgaatgaga aaaactgatg ctggtcaagg tgcaagtgac 840
attgatccaa tgagtgttga ttcgtcggtt ggtttcgatc aagttggagg actcggtcat 900
catattcaat ctctcaaaga agttgtcctt ttcccaatgc tttatcctga agtatttgaa 960
aagttcagaa taaatccgcc aaaaggtgtc gtattctatg ggcccccagg aaccggaaaa 1020
actcttgttg cacgtgctct tgcgaatgaa tgccgaagag gagccaacaa agttgcattc 1080
tttatgcgaa aaggagctga ttgtttgagt aaatgggtcg gagagtctga aagacaactt 1140
cgactattgt ttgatcaagc atatgccatg cgtccatcaa ttatcttttt tgatgaaatc 1200
gatggtttag ctccagttcg ttcttcgaaa caagatcaaa ttcatgcgag catagtgagt 1260
acattgttag cactgatgga cggattagat ggtcgaggag aagttgttgt aattggagct 1320
acaaatcgat tagatactct tgatccagct ctacgaagac caggacgttt tgatagggaa 1380
ctccgttttt cccttcctga tttgaatgcc cgacgacaaa ttttagatat tcacacttcc 1440
aaatgggagg aaaacaaacc gattccggaa acactcgatg ctattgcaga gagaactagt 1500
ggatattgcg gtgccgattt gaaattcttg tgcacagaag cagttcttat tggattgaga 1560
tctcgatatc ctcatatcta catgtgttct gaaaggctga aactcgacgt tgccaccatt 1620
aaaatcactt ctgagcattt tggacatgca atgcggagga tcacacctgc ttccagacga 1680
gatcttacaa ttccatcgag accattggat gaacgaactt caattttact tggagatact 1740
gtgagcaact tgatttcgtt gagaattcca caaggatatc gatgcgtgga aaatgcaatg 1800
gcgacagctt caagcgaact ggaacaagtt gttagggcat tggaacctaa tccaaccgtg 1860
cctgctattc gtcttcttct atgtggatct gagcaactcg cggacggtgg acaaacgtct 1920
tacgttcttc ccgccattct agcaaaactg gatcatctac cggtgttctc tctttctgtt 1980
tcaagtcttt taacagatgg acgtccagaa gaagcatttt caaatgcaat tcagtcggca 2040
atgcgtgcat cagcaaccgg cccatgtatc atgcttcttc catcaatcga tgaatggatt 2100
aaagtcattc ctgtatcagt tcaacatatg cttatcacat gtctggaatc aatgactggt 2160
ttcacgccga ttcttttcct atcgacgctt gacacttcgt ttgaagatgc gcctgaatat 2220
gtcacggaaa ttttcagaca tgcaaactgt atcactttga atccttcccg ccgaacaatt 2280
cgtcaaaaat atttcgagca cgttattgag aagattaaca ctccaccaaa agtatttgat 2340
ccacgtctta tgcgagatcg tcgattcgtt gaatttgtcg agcccgtgga tccggatgag 2400
gctgaggact attatgagat tatcgaaaca ccgatttgta tgcaggatat tatggaaaag 2460
ctgaataatt gcgaatataa tcatgcggac aaatttgttg ctgacttgat attgattcaa 2520
acgaatgctc tcgaatataa tccgtcgaca acaaaagacg gaaaactgat tcgtcaaatg 2580
gcaaatacac ttcgagatgc aattgatgat ctgatcgagt gtgaacttga cgaatcgttt 2640
gtcgaacgaa ttgaaacagt ttcacgaatg ctacaagacg ctggagtaac tccaacctct 2700
gataaattac tcacagaaat tcccaaaggt tttgcccgaa aaaaggcgtg gtctatgacc 2760
aattcacttg caaaagaaat tgagcaatgg acatcagagc gtgaagctga aaatcaaaaa 2820
atgttgagca agcttggtgt agctgctcca acattggaac tcgttgttgt ccctgtagaa 2880
gatatgaaaa gtgaagaagg aacgagcaca tcaactgatg gagttccagc ttctgctgga 2940
aataaaaaga aattgctcaa aaagaagaag ggtcaaaaga aaagcaaaac cggcgaatct 3000

gaagagcacg atgaagatag tacagttgaa gacgcaggag aagatacaat tgtagaaaat 3060
cttgaaatca agaaaaatca ggaaactcca aacagtgaac atgatatcga gatgaaagat 3120
gcttctaaag attctacacc ttccgtgcaa atttcgattg ctgagaaaga attgatagtt 3180
tcgaagccag caacatgcga gctgattcag tgttgtgtcg aaaagtcgga gggttggtca 3240
gtcagcgagt tggaacgttt gtcttctgtt ctatcacata caattgaacg attccgggac 3300
gaatggaacc gcgaaaatct accagcccaa cttacccaaa ttgtacgtga atggcaaaca 3360
gctgacgatt cgaacaatac aattgttaat ggaaccctca acaaatcgaa tggaaatctt 3420
gccaatggac actaa 3435

<210> 24
<211> 1144
<212> PRT
<213> Caenorhabditis elegans

<400> 24
Met Gln Ala Ile Asp Glu Ser Lys Arg Asn Gln Lys Val Pro Pro Ala
1 5 10 15

Lys Arg Lys Arg Ile Tyr Leu Ser Asp Glu Glu Glu Glu Asp Phe Ala
20 25 30

Glu Ala Ala His Val Glu Asn Thr Val Pro Glu Arg Ala Thr Arg Arg
35 40 45

Ser Thr Arg Arg Arg Ser Ser Met His Glu Glu Leu Gly Val Ser Glu
50 55 60

Gln Glu Glu Ser Pro Val Arg Arg Thr Arg Lys Ala Ala Lys Arg Leu
65 70 75 80

Gly Ser Glu Gln Pro Glu Glu Asn Leu Ala Ala Asp Asp Pro Leu Pro
85 90 95

Met Glu Gly Gly Gly Glu Ile Val Leu Pro Ile Ala Glu Ile Asp Gly
100 105 110

Met Ala Glu Gln Glu Asn Glu Asp Leu Ile Glu Lys Ile Gly Arg Glu
115 120 125

Glu Glu Glu Glu Gly Ala Glu Glu Asp Glu Gln Ser Gly Glu Lys Asp
130 135 140

Pro Glu Glu Glu Glu Asp Asp Ser Ser Asn Ala Glu Ser Ser Glu Glu
145 150 155 160

Ser Thr Ala Pro Arg Gln Tyr Ser Leu Arg Arg Arg Gln Pro Val Val
165 170 175

Gln Phe Asn Ala Ser Glu Ala Arg Glu Asn Arg Arg Ala Arg Leu Glu
180 185 190

His His Arg Val Ala Asn Gln Asn Arg His His Arg Asn Arg Asn Gly
195 200 205

Ser Arg Arg Arg Arg Ser Asp Ser Asp Ser Asp Ser Asp Asp Met Val
210 215 220

Leu Pro Arg Pro Asp Lys Arg Gln Ser Arg Pro His Met His Asn Arg

```
225                 230                 235                 240

Gly Glu Arg Glu Arg Gly Arg Phe Met Pro Ile Asn Met Thr Glu Lys
                245                 250                 255

Glu Leu Gln Ser Ala Gln His Ile Leu Met Asp Arg Met Arg Lys Thr
            260                 265                 270

Asp Ala Gly Gln Gly Ala Ser Asp Ile Asp Pro Met Ser Val Asp Ser
        275                 280                 285

Ser Val Gly Phe Asp Gln Val Gly Gly Leu Gly His His Ile Gln Ser
    290                 295                 300

Leu Lys Glu Val Val Leu Phe Pro Met Leu Tyr Pro Glu Val Phe Glu
305                 310                 315                 320

Lys Phe Arg Ile Asn Pro Pro Lys Gly Val Val Phe Tyr Gly Pro Pro
                325                 330                 335

Gly Thr Gly Lys Thr Leu Val Ala Arg Ala Leu Ala Asn Glu Cys Arg
            340                 345                 350

Arg Gly Ala Asn Lys Val Ala Phe Phe Met Arg Lys Gly Ala Asp Cys
        355                 360                 365

Leu Ser Lys Trp Val Gly Glu Ser Glu Arg Gln Leu Arg Leu Leu Phe
    370                 375                 380

Asp Gln Ala Tyr Ala Met Arg Pro Ser Ile Ile Phe Phe Asp Glu Ile
385                 390                 395                 400

Asp Gly Leu Ala Pro Val Arg Ser Ser Lys Gln Asp Gln Ile His Ala
                405                 410                 415

Ser Ile Val Ser Thr Leu Leu Ala Leu Met Asp Gly Leu Asp Gly Arg
            420                 425                 430

Gly Glu Val Val Val Ile Gly Ala Thr Asn Arg Leu Asp Thr Leu Asp
        435                 440                 445

Pro Ala Leu Arg Arg Pro Gly Arg Phe Asp Arg Glu Leu Arg Phe Ser
    450                 455                 460

Leu Pro Asp Leu Asn Ala Arg Arg Gln Ile Leu Asp Ile His Thr Ser
465                 470                 475                 480

Lys Trp Glu Glu Asn Lys Pro Ile Pro Glu Thr Leu Asp Ala Ile Ala
                485                 490                 495

Glu Arg Thr Ser Gly Tyr Cys Gly Ala Asp Leu Lys Phe Leu Cys Thr
            500                 505                 510

Glu Ala Val Leu Ile Gly Leu Arg Ser Arg Tyr Pro His Ile Tyr Met
        515                 520                 525

Cys Ser Glu Arg Leu Lys Leu Asp Val Ala Thr Ile Lys Ile Thr Ser
    530                 535                 540
```

Glu His Phe Gly His Ala Met Arg Arg Ile Thr Pro Ala Ser Arg Arg
545 550 555 560

Asp Leu Thr Ile Pro Ser Arg Pro Leu Asp Glu Arg Thr Ser Ile Leu
565 570 575

Leu Gly Asp Thr Val Ser Asn Leu Ile Ser Leu Arg Ile Pro Gln Gly
580 585 590

Tyr Arg Cys Val Glu Asn Ala Met Ala Thr Ala Ser Ser Glu Leu Glu
595 600 605

Gln Val Val Arg Ala Leu Glu Pro Asn Pro Thr Val Pro Ala Ile Arg
610 615 620

Leu Leu Leu Cys Gly Ser Glu Gln Leu Ala Asp Gly Gly Gln Thr Ser
625 630 635 640

Tyr Val Leu Pro Ala Ile Leu Ala Lys Leu Asp His Leu Pro Val Phe
645 650 655

Ser Leu Ser Val Ser Ser Leu Leu Thr Asp Gly Arg Pro Glu Glu Ala
660 665 670

Phe Ser Asn Ala Ile Gln Ser Ala Met Arg Ala Ser Ala Thr Gly Pro
675 680 685

Cys Ile Met Leu Leu Pro Ser Ile Asp Glu Trp Ile Lys Val Ile Pro
690 695 700

Val Ser Val Gln His Met Leu Ile Thr Cys Leu Glu Ser Met Thr Gly
705 710 715 720

Phe Thr Pro Ile Leu Phe Leu Ser Thr Leu Asp Thr Ser Phe Glu Asp
725 730 735

Ala Pro Glu Tyr Val Thr Glu Ile Phe Arg His Ala Asn Cys Ile Thr
740 745 750

Leu Asn Pro Ser Arg Arg Thr Ile Arg Gln Lys Tyr Phe Glu His Val
755 760 765

Ile Glu Lys Ile Asn Thr Pro Pro Lys Val Phe Asp Pro Arg Leu Met
770 775 780

Arg Asp Arg Arg Phe Val Glu Phe Val Glu Pro Val Asp Pro Asp Glu
785 790 795 800

Ala Glu Asp Tyr Tyr Glu Ile Ile Glu Thr Pro Ile Cys Met Gln Asp
805 810 815

Ile Met Glu Lys Leu Asn Asn Cys Glu Tyr Asn His Ala Asp Lys Phe
820 825 830

Val Ala Asp Leu Ile Leu Ile Gln Thr Asn Ala Leu Glu Tyr Asn Pro
835 840 845

Ser Thr Thr Lys Asp Gly Lys Leu Ile Arg Gln Met Ala Asn Thr Leu
850 855 860

Arg Asp Ala Ile Asp Asp Leu Ile Glu Cys Glu Leu Asp Glu Ser Phe
865 870 875 880

Val Glu Arg Ile Glu Thr Val Ser Arg Met Leu Gln Asp Ala Gly Val
885 890 895

Thr Pro Thr Ser Asp Lys Leu Leu Thr Glu Ile Pro Lys Gly Phe Ala
900 905 910

Arg Lys Lys Ala Trp Ser Met Thr Asn Ser Leu Ala Lys Glu Ile Glu
915 920 925

Gln Trp Thr Ser Glu Arg Glu Ala Glu Asn Gln Lys Met Leu Ser Lys
930 935 940

Leu Gly Val Ala Ala Pro Thr Leu Glu Leu Val Val Val Pro Val Glu
945 950 955 960

Asp Met Lys Ser Glu Glu Gly Thr Ser Thr Ser Thr Asp Gly Val Pro
965 970 975

Ala Ser Ala Gly Asn Lys Lys Lys Leu Leu Lys Lys Lys Lys Gly Gln
980 985 990

Lys Lys Ser Lys Thr Gly Glu Ser Glu Glu His Asp Glu Asp Ser Thr
995 1000 1005

Val Glu Asp Ala Gly Glu Asp Thr Ile Val Glu Asn Leu Glu Ile Lys
1010 1015 1020

Lys Asn Gln Glu Thr Pro Asn Ser Glu His Asp Ile Glu Met Lys Asp
1025 1030 1035 1040

Ala Ser Lys Asp Ser Thr Pro Ser Val Gln Ile Ser Ile Ala Glu Lys
1045 1050 1055

Glu Leu Ile Val Ser Lys Pro Ala Thr Cys Glu Leu Ile Gln Cys Cys
1060 1065 1070

Val Glu Lys Ser Glu Gly Trp Ser Val Ser Glu Leu Glu Arg Leu Ser
1075 1080 1085

Ser Val Leu Ser His Thr Ile Glu Arg Phe Arg Asp Glu Trp Asn Arg
1090 1095 1100

Glu Asn Leu Pro Ala Gln Leu Thr Gln Ile Val Arg Glu Trp Gln Thr
1105 1110 1115 1120

Ala Asp Asp Ser Asn Asn Thr Ile Val Asn Gly Thr Leu Asn Lys Ser
1125 1130 1135

Asn Gly Asn Leu Ala Asn Gly His
1140

<210> 25
<211> 1908
<212> DNA
<213> Caenorhabditis elegans

<400> 25
gaagaagaag aggatgatag cagcaatgca gaaagttcag aggaatcaac tgctccacgt 60
caatacagtc ttcgccgccg tcaaccagtt gtccagttca atgcatcaga agctcgtgaa 120
aatcgacggg ctcgtcttga acatcatcgg gttgcaaatc aaaacagaca tcaccgcaat 180
cgaaatggat ctcgtcgaag acgaagcgat tcggattctg atagtgatga tatggtgctt 240
ccgagacctg ataaacgtca atcgaggcct catatgcata accgtggaga acgtgaacgt 300
ggtcgattta tgccgatcaa tatgaccgaa aaagaattgc agtctgccca acacattctg 360
atggatcgaa tgagaaaaac tgatgctggt caaggtgcaa gtgacattga tccaatgagt 420
gttgattcgt cggttggttt cgatcaagtt ggaggactcg gtcatcatat tcaatctctc 480
aaagaagttg tcctttcccc aatgctttat cctgaagtat ttgaaaagtt cagaataaat 540
ccgccaaaag gtgtcgtatt ctatgggccc ccaggaaccg gaaaaactct tgttgcacgt 600
gctcttgcga atgaatgccg aagaggagcc aacaaagttg cattctttat gcgaaaagga 660
gctgattgtt tgagtaaatg ggtcggagag tctgaaagac aacttcgact attgtttgat 720
caagcatatg ccatgcgtcc atcaattatc tttttgatg aaatcgatgg tttagctcca 780
gttcgttctt cgaaacaaga tcaaattcat gcgagcatag tgagtacatt gttagcactg 840
atggacggat tagatggtcg aggagaagtt gttgtaattg gagctacaaa tcgattagat 900
actcttgatc cagctctacg aagaccagga cgttttgata gggaactccg tttttccctt 960
cctgatttga atgcccgacg acaaatttta gatattcaca cttccaaatg ggaggaaaac 1020
aaaccgattc cggaaacact cgatgctatt gcagagagaa ctagtggata ttgcggtgcc 1080
gatttgaaat tcttgtgcac agaagcagtt cttattggat tgagatctcg atatcctcat 1140
atctacatgt gttctgaaag gctgaaactc gacgttgcca ccattaaaat cacttctgag 1200
cattttggac atgcaatgcg gaggatcaca cctgcttcca gacgagatct tacaattcca 1260
tcgagaccat tggatgaacg aacttcaatt ttacttggag atactgtgag caacttgatt 1320
tcgttgagaa ttccacaagg atatcgatgc gtggaaaatg caatggcgac agcttcaagc 1380
gaactggaac aagttgttag ggcattggaa cctaatccaa ccgtgcctgc tattcgtctt 1440
cttctatgtg gatctgagca actcgcggac ggtggacaaa cgtcttacgt tcttcccgcc 1500
attctagcaa aactggatca tctaccggtg ttctctcttt ctgtttcaag tcttttaaca 1560
gatggacgtc cagaagaagc attttcaaat gcaattcagt cggcaatgcg tgcatcagca 1620
accggcccat gtatcatgct tcttccatca atcgatgaat ggattaaagt cattcctgta 1680
tcagttcaac atatgcttat cacatgtctg gaatcaatga ctggtttcac gccgattctt 1740
ttcctatcga cgcttgacac ttcgtttgaa gatgcgcctg aatatgtcac ggaaatttc 1800
agacatgcaa actgtatcac tttgaatcct tcccgccgaa caattcgtca aaaatatttc 1860
gagcacgtta ttgagaagat taacactcca ccaaaagtat ttgatcca 1908

<210> 26
<211> 636
<212> PRT
<213> Caenorhabditis elegans

<400> 26
Glu Glu Glu Glu Asp Asp Ser Ser Asn Ala Glu Ser Ser Glu Glu Ser
1 5 10 15

Thr Ala Pro Arg Gln Tyr Ser Leu Arg Arg Arg Gln Pro Val Val Gln
20 25 30

Phe Asn Ala Ser Glu Ala Arg Glu Asn Arg Arg Ala Arg Leu Glu His
35 40 45

His Arg Val Ala Asn Gln Asn Arg His His Arg Asn Arg Asn Gly Ser
50 55 60

Arg Arg Arg Arg Ser Asp Ser Asp Ser Asp Ser Asp Asp Met Val Leu

65 70 75 80

Pro Arg Pro Asp Lys Arg Gln Ser Arg Pro His Met His Asn Arg Gly
85 90 95

Glu Arg Glu Arg Gly Arg Phe Met Pro Ile Asn Met Thr Glu Lys Glu
100 105 110

[illegible] [illegible] Ser Ala Gln His Ile Leu Met Asp Arg Met Arg Lys Thr Asp
115 120 125

[illegible] [illegible] Gln Gly Ala Ser Asp Ile Asp Pro Met Ser Val Asp Ser Ser
[illegible] 135 140

[illegible] [illegible] Phe Asp Gln Val Gly Gly Leu Gly His His Ile Gln Ser Leu
[illegible] 150 155 160

Tyr [illegible] Val Val Leu Phe Pro Met Leu Tyr Pro Glu Val Phe Glu Lys
165 170 175

Phe Arg Ile Asn Pro Pro Lys Gly Val Val Phe Tyr Gly Pro Pro Gly
180 185 190

Thr Gly Lys Thr Leu Val Ala Arg Ala Leu Ala Asn Glu Cys Arg Arg
195 200 205

Gly Ala Asn Lys Val Ala Phe Phe Met Arg Lys Gly Ala Asp Cys Leu
210 215 220

Ser Lys Trp Val Gly Glu Ser Glu Arg Gln Leu Arg Leu Leu Phe Asp
225 230 235 240

Gln Ala Tyr Ala Met Arg Pro Ser Ile Ile Phe Phe Asp Glu Ile Asp
245 250 255

Gly Leu Ala Pro Val Arg Ser Ser Lys Gln Asp Gln Ile His Ala Ser
260 265 270

Ile Val Ser Thr Leu Leu Ala Leu Met Asp Gly Leu Asp Gly Arg Gly
275 280 285

Glu Val Val Val Ile Gly Ala Thr Asn Arg Leu Asp Thr Leu Asp Pro
290 295 300

Ala Leu Arg Arg Pro Gly Arg Phe Asp Arg Glu Leu Arg Phe Ser Leu
305 310 315 320

Pro Asp Leu Asn Ala Arg Arg Gln Ile Leu Asp Ile His Thr Ser Lys
325 330 335

Trp Glu Glu Asn Lys Pro Ile Pro Glu Thr Leu Asp Ala Ile Ala Glu
340 345 350

Arg Thr Ser Gly Tyr Cys Gly Ala Asp Leu Lys Phe Leu Cys Thr Glu
355 360 365

Ala Val Leu Ile Gly Leu Arg Ser Arg Tyr Pro His Ile Tyr Met Cys
370 375 380

Ser Glu Arg Leu Lys Leu Asp Val Ala Thr Ile Lys Ile Thr Ser Glu
385 390 395 400

His Phe Gly His Ala Met Arg Arg Ile Thr Pro Ala Ser Arg Arg Asp
405 410 415

Leu Thr Ile Pro Ser Arg Pro Leu Asp Glu Arg Thr Ser Ile Leu Leu
420 425 430

Gly Asp Thr Val Ser Asn Leu Ile Ser Leu Arg Ile Pro Gln Gly Tyr
435 440 445

Arg Cys Val Glu Asn Ala Met Ala Thr Ala Ser Ser Glu Leu Glu Gln
450 455 460

Val Val Arg Ala Leu Glu Pro Asn Pro Thr Val Pro Ala Ile Arg Leu
465 470 475 480

Leu Leu Cys Gly Ser Glu Gln Leu Ala Asp Gly Gly Gln Thr Ser Tyr
485 490 495

Val Leu Pro Ala Ile Leu Ala Lys Leu Asp His Leu Pro Val Phe Ser
500 505 510

Leu Ser Val Ser Ser Leu Leu Thr Asp Gly Arg Pro Glu Glu Ala Phe
515 520 525

Ser Asn Ala Ile Gln Ser Ala Met Arg Ala Ser Ala Thr Gly Pro Cys
530 535 540

Ile Met Leu Leu Pro Ser Ile Asp Glu Trp Ile Lys Val Ile Pro Val
545 550 555 560

Ser Val Gln His Met Leu Ile Thr Cys Leu Glu Ser Met Thr Gly Phe
565 570 575

Thr Pro Ile Leu Phe Leu Ser Thr Leu Asp Thr Ser Phe Glu Asp Ala
580 585 590

Pro Glu Tyr Val Thr Glu Ile Phe Arg His Ala Asn Cys Ile Thr Leu
595 600 605

Asn Pro Ser Arg Arg Thr Ile Arg Gln Lys Tyr Phe Glu His Val Ile
610 615 620

Glu Lys Ile Asn Thr Pro Pro Lys Val Phe Asp Pro
625 630 635

<210> 27
<211> 3024
<212> DNA
<213> Caenorhabditis elegans

<400> 27
atggaacgac acggttcaga attacaagga gaagtggttt tcactacgac gagtactgcg 60
aataacaatg tgaaacgagc cgtcgaatat ttcgaagatg ataatcaaga tgatgcacta 120

acatccactt catccggcaa cagtacacag aaggaatcgt ctccgttcac agattttgat 180
gatgttccgc cgccaccagt ggctccagaa actccggctc cagcccaaaa ccgccgagaa 240
tctgcgtccc ccgaacgtca attccttgat gagagtcacc tcggcggaat tggatctccg 300
ttatcacaga gtacaaggct cgatgagacg tttattgagg aatactctat tgagttggac 360
acctctggga aaaacaacat ttcctccgca gcttcacctg gtcccaagtc tccatttgac 420
gatgatttta ctgatacagc tgccccggta gccccgccac cagcacctac aaaagctgcc 480
gaggaatatc gccggcaacc acatcagaat ccgtttgatg aggaagagga agaagaatca 540
cagtttggag gggaactctt gtccggtaga gacccatttg atgaagattc tggaaactct 600
aacgaaaacc agctccgcga gaagaagctc cacaaaaaag agcaactagc tcaccggctc 660
tcctcctctt ccgaagaaat tgtggaagcc tcaattcacg aggatgagcc aattgtgatg 720
gctcaaattc cagaggagaa gcccaaaccg aaggctatcc cggcttttga taacgcctat 780
gacgccgact ttgacaattc cccaccactt catcattact cggcggttca tttggaaact 840
gggctctcac cgcttgagga ggctcagaga gctctccgag ccaaccgagc aaggcataag 900
ccgtcgaatg tatcgttggc agaggaggcg aagcttgctg ccagacagag atactcgaat 960
gcatcggata ttagaagaga agaggaggaa gaagtggtgg aagaagatcc agcagtggta 1020
gttccagttc tacggaagga tctggaagtt gaggaggctc cgaaatcggt acgaccgcct 1080
agatatcgga agtcaaggga aattgaggaa ccggtggttg ttgatcggtt tgttgaggaa 1140
gaagttgatg agaaggagga tattgatgcg atttttgaga aatatcgaaa gacgtctgta 1200
tccgccgatc caaaatctca cacgccgatt ttgatggccg acgagtacaa agaaccacag 1260
aagcaagtgc ctgcaccagt tgtcgttgct caggagtcgc cgatcttaaa acgaagaaac 1320
tctctagtac catcaaggat cagtggaagg cagtcgacgc ggcgaagtgt caccagtgtt 1380
cggagtatga gaggaaaacg gaagacccgt gcgattcccg agttcttcga cctaacccga 1440
catcagaaca ttcggctccg tgctccagct acgaaaaaga aacgcataag cctacaccgc 1500
gtcgaggaca ccgaagtcgt cgtcgagctc ctgaacggcc aaaaagttga agtggcatgc 1560
cgaagtgacg tcatctcgcg tgacgtgttc tctcttatcg tccagaacat gaatatcaat 1620
gagcacgtgt tcttcgggct ctcgttcctt cgtgatggcg agcactactt tatagaggat 1680
catcaacggc tcgaaaagtt tgcaccgtcc ggatggaaaa gtgtggcacg cgtgggcgtt 1740
aaggtgccat acgtgttgca cttacgattc aaattttatc cgcagatttt ggatttcatc 1800
aaaacagacg taacgatgaa cgagttgtat ctacaatgta gacgagacgt actcgaagag 1860
agaatccaac ccaaaagaga cgcagcattc gaactcgccg cgctcgccct ccaagccgaa 1920
tttggaaatc gaccgccacc ggtgatcacc gattatttcg atattcaaca ctacctgccc 1980
aagaagtata gctcatttga agatcagtca cggctgaaga acattctcgc cgagctgcac 2040
gggcattacg cgggcacccg gatttctgag gccaagcaca aatacattca aatctgtcag 2100
cgacacccag attttggagc gcatgtgcat cgtgtgttcc gaacgaagcc aacatcagct 2160
cacggagcat ctccgttcga tcctgatact gggtcatcgc tttggattgg aattatgcca 2220
cgtggaatat cgatttatga gcagcaggga ggagcccgtg aggtgattgc cgagcatgtt 2280
tggccgcaga cgcagaccct gcaattcgat aagaagcgat ttgtgattgt tgccgttgga 2340
gctcacgatg agcaaattga gagcacattt tatacggatc atcactcgaa atcctcctat 2400
ttcgtgcgat tcgccgcctc ccaacatcgt tggatgatga agatgcgtca gtggaagagc 2460
acgctccggc acgagaatac gattcaggcg atgccagacg tgattgttga agggcagact 2520
atcccgccgg cgcctatacg gcaaggtctc gaagaatcac ctccatcaac tccactgctc 2580
gcctctgccg atcagatctt tgcgaaaatg tcagtgagcc aggaaaaacc ggcggccaac 2640
cgacccgccg aactgccacc accggcaccg tcgtcgaaat tcgcggctca atatgatacg 2700
gttgatgaag gaattgtttg tgattcccaa gctgaaaact ttgagcgcgt cgactcgact 2760
gacaacggga acgtgacgcc acgtggcatg caatttgaca ttttgctcgt gaaagacccc 2820
gccaacggtc tgggtcttac cctagtcgac gggaatctca acggagttcc aggagtgtat 2880
gtgaagcttg tcgccgacaa tggagccgga atgaaagcgg taagaattcg aaatttcagt 2940
caatatccct tttcctcggg gtgtacacta gagttgtgta aatctacaca taacgttttt 3000
tcgattatct ctgaaaaatc gtaa 3024

<210> 28
<211> 1007
<212> PRT
<213> Caenorhabditis elegans

<400> 28
Met Glu Arg His Gly Ser Glu Leu Gln Gly Glu Val Val Phe Thr Thr
1 5 10 15

Thr Ser Thr Ala Asn Asn Asn Val Lys Arg Ala Val Glu Tyr Phe Glu
20 25 30

Asp Asp Asn Gln Asp Asp Ala Leu Thr Ser Thr Ser Ser Gly Asn Ser
35 40 45

Thr Gln Lys Glu Ser Ser Pro Phe Thr Asp Phe Asp Asp Val Pro Pro
50 55 60

[illegible] [illegible] Val Ala Pro Glu Thr Pro Ala Pro Ala Gln Asn Arg Arg Glu
[illegible] 70 75 80

[illegible] Ala Ser Pro Glu Arg Gln Phe Leu Asp Glu Ser His Leu Gly Gly
85 90 95

[illegible] [illegible] Ser Pro Leu Ser Gln Ser Thr Arg Leu Asp Glu Thr Phe Ile
100 105 110

[illegible] Glu Tyr Ser Ile Glu Leu Asp Thr Ser Gly Lys Asn Asn Ile Ser
115 120 125

[illegible] Ala Ala Ser Pro Gly Pro Lys Ser Pro Phe Asp Asp Asp Phe Thr
130 135 140

Asp Thr Ala Ala Pro Val Ala Pro Pro Pro Ala Pro Thr Lys Ala Ala
145 150 155 160

Glu Glu Tyr Arg Arg Gln Pro His Gln Asn Pro Phe Asp Glu Glu Glu
165 170 175

Glu Glu Glu Ser Gln Phe Gly Gly Gly Thr Leu Ser Gly Arg Asp Pro
180 185 190

Phe Asp Glu Asp Ser Gly Asn Ser Asn Glu Asn Gln Leu Arg Glu Lys
195 200 205

Lys Leu His Lys Lys Glu Gln Leu Ala His Arg Leu Ser Ser Ser Ser
210 215 220

Glu Glu Ile Val Glu Ala Ser Ile His Glu Asp Glu Pro Ile Val Met
225 230 235 240

Ala Gln Ile Pro Glu Glu Lys Pro Lys Pro Lys Ala Ile Pro Ala Phe
245 250 255

Asp Asn Ala Tyr Asp Ala Asp Phe Asp Asn Ser Pro Pro Leu His His
260 265 270

Tyr Ser Ala Val His Leu Glu Thr Gly Leu Ser Pro Leu Glu Glu Ala
275 280 285

Gln Arg Ala Leu Arg Ala Asn Arg Ala Arg His Lys Pro Ser Asn Val
290 295 300

Ser Leu Ala Glu Glu Ala Lys Leu Ala Ala Arg Gln Arg Tyr Ser Asn
305 310 315 320

Ala Ser Asp Ile Arg Arg Glu Glu Glu Glu Glu Val Val Glu Glu Asp
325 330 335

Pro Ala Val Val Val Pro Val Leu Arg Lys Asp Leu Glu Val Glu Glu
340 345 350

Ala Pro Lys Ser Val Arg Pro Pro Arg Tyr Arg Lys Ser Arg Glu Ile
355 360 365

Glu Glu Pro Val Val Val Asp Arg Phe Val Glu Glu Glu Val Asp Glu
370 375 380

Lys Glu Asp Ile Asp Ala Ile Phe Glu Lys Tyr Arg Lys Thr Ser Val
385 390 395 400

Ser Ala Asp Pro Lys Ser His Thr Pro Ile Leu Met Ala Asp Glu Tyr
405 410 415

Lys Glu Pro Gln Lys Gln Val Pro Ala Pro Val Val Val Ala Gln Glu
420 425 430

Ser Pro Ile Leu Lys Arg Arg Asn Ser Leu Val Pro Ser Arg Ile Ser
435 440 445

Gly Arg Gln Ser Thr Arg Arg Ser Val Thr Ser Val Arg Ser Met Arg
450 455 460

Gly Lys Arg Lys Thr Arg Ala Ile Pro Glu Phe Phe Asp Leu Thr Arg
465 470 475 480

His Gln Asn Ile Arg Leu Arg Ala Pro Ala Thr Lys Lys Lys Arg Ile
485 490 495

Ser Leu His Arg Val Glu Asp Thr Glu Val Val Val Glu Leu Leu Asn
500 505 510

Gly Gln Lys Val Glu Val Ala Cys Arg Ser Asp Val Ile Ser Arg Asp
515 520 525

Val Phe Ser Leu Ile Val Gln Asn Met Asn Ile Asn Glu His Val Phe
530 535 540

Phe Gly Leu Ser Phe Leu Arg Asp Gly Glu His Tyr Phe Ile Glu Asp
545 550 555 560

His Gln Arg Leu Glu Lys Phe Ala Pro Ser Gly Trp Lys Ser Val Ala
565 570 575

Arg Val Gly Val Lys Val Pro Tyr Val Leu His Leu Arg Phe Lys Phe
580 585 590

Tyr Pro Gln Ile Leu Asp Phe Ile Lys Thr Asp Val Thr Met Asn Glu
595 600 605

Leu Tyr Leu Gln Cys Arg Arg Asp Val Leu Glu Glu Arg Ile Gln Pro
610 615 620

Lys Arg Asp Ala Ala Phe Glu Leu Ala Ala Leu Ala Leu Gln Ala Glu

625 630 635 640

Phe Gly Asn Arg Pro Pro Pro Val Ile Thr Asp Tyr Phe Asp Ile Gln
645 650 655

His Tyr Leu Pro Lys Lys Tyr Ser Ser Phe Glu Asp Gln Ser Arg Leu
660 665 670

Lys Asn Ile Leu Ala Glu Leu His Gly His Tyr Ala Gly Thr Arg Ile
675 680 685

Ser Glu Ala Lys His Lys Tyr Ile Gln Ile Cys Gln Arg His Pro Asp
690 695 700

Phe Gly Ala His Val His Arg Val Phe Arg Thr Lys Pro Thr Ser Ala
705 710 715 720

His Gly Ala Ser Pro Phe Asp Pro Asp Thr Gly Ser Ser Leu Trp Ile
725 730 735

Gly Ile Met Pro Arg Gly Ile Ser Ile Tyr Glu Gln Gln Gly Gly Ala
740 745 750

Arg Glu Val Ile Ala Glu His Val Trp Pro Gln Thr Gln Thr Leu Gln
755 760 765

Phe Asp Lys Lys Arg Phe Val Ile Val Ala Val Gly Ala His Asp Glu
770 775 780

Gln Ile Glu Ser Thr Phe Tyr Thr Asp His His Ser Lys Ser Ser Tyr
785 790 795 800

Phe Val Arg Phe Ala Ala Ser Gln His Arg Trp Met Met Lys Met Arg
805 810 815

Gln Trp Lys Ser Thr Leu Arg His Glu Asn Thr Ile Gln Ala Met Pro
820 825 830

Asp Val Ile Val Glu Gly Gln Thr Ile Pro Pro Ala Pro Ile Arg Gln
835 840 845

Gly Leu Glu Glu Ser Pro Pro Ser Thr Pro Leu Leu Ala Ser Ala Asp
850 855 860

Gln Ile Phe Ala Lys Met Ser Val Ser Gln Glu Lys Pro Ala Ala Asn
865 870 875 880

Arg Pro Ala Glu Leu Pro Pro Pro Ala Pro Ser Ser Lys Phe Ala Ala
885 890 895

Gln Tyr Asp Thr Val Asp Glu Gly Ile Val Cys Asp Ser Gln Ala Glu
900 905 910

Asn Phe Glu Arg Val Asp Ser Thr Asp Asn Gly Asn Val Thr Pro Arg
915 920 925

Gly Met Gln Phe Asp Ile Leu Leu Val Lys Asp Pro Ala Asn Gly Leu
930 935 940

```
Gly Leu Thr Leu Val Asp Gly Asn Leu Asn Gly Val Pro Gly Val Tyr
945                 950                 955                 960

Val Lys Leu Val Ala Asp Asn Gly Ala Gly Met Lys Ala Val Arg Ile
                965                 970                 975

Arg Asn Phe Ser Gln Tyr Pro Phe Ser Ser Gly Cys Thr Leu Glu Leu
            980                 985                 990

Cys Lys Ser Thr His Asn Val Phe Ser Ile Ile Ser Glu Lys Ser
        995                 1000                1005


<210> 29
<211> 1311
<212> DNA
<213> Caenorhabditis elegans

<400> 29
aaaccgaagg ctatcccggc ttttgataac gcctatgacg ccgactttga caattcccca 60
ccacttcatc attactcggc ggttcatttg gaaactgggc tctcaccgct tgaggaggct 120
cagagagctc tccgagccaa ccgagcaagg cataagccgt cgaatgtatc gttggcagag 180
gaggcgaagc ttgctgccag acagagatac tcgaatgcat cggatattag aagagaagag 240
gaggaagaag tggtggaaga agatccagca gtggtagttc cagttctacg gaaggatctg 300
gaagttgagg aggctccgaa atcggtacga ccgcctagat atcggaagtc aagggaaatt 360
gaggaaccgg tggttgttga tcggtttgtt gaggaagaag ttgatgagaa ggaggatatt 420
gatgcgattt ttgagaaata tcgaaagacg tctgtatccg ccgatccaaa atctcacacg 480
ccgattttga tggccgacga gtacaaagaa ccacagaagc aagtgcctgc accagttgtc 540
gttgctcagg agtcgccgat cttaaaacga agaaactctc tagtaccatc aaggatcagt 600
ggaaggcagt cgacgcggcg aagtgtcacc agtgttcgga gtatgagagg aaaacggaag 660
acccgtgcga ttcccgagtt cttcgaccta acccgacatc agaacattcg gctccgtgct 720
ccagctacga aaaagaaacg cataagccta caccgcgtcg aggacaccga agtcgtcgtc 780
gagctcctga acggccaaaa agttgaagtg gcatgccgaa gtgacgtcat ctcgcgtgac 840
gtgttctctc ttatcgtcca gaacatgaat atcaatgagc acgtgttctt cgggctctcg 900
ttccttcgtg atggcgagca ctactttata gaggatcatc aacggctcga aaagtttgca 960
ccgtccggat ggaaaagtgt ggcacgcgtg ggcgttaagg tgccatacgt gttgcactta 1020
cgattcaaat tttatccgca gattttggat ttcatcaaaa cagacgtaac gatgaacgag 1080
ttgtatctac aatgtagacg agacgtactc gaagagagaa tccaacccaa aagagacgca 1140
gcattcgaac tcgccgcgct cgccctccaa gccgaatttg gaaatcgacc gccaccggtg 1200
atcaccgatt atttcgatat tcaacactac ctgcccaaga agtatagctc atttgaagat 1260
cagtcacggc tgaagaacat tctcgccgag ctgcacgggc attacgcggg c          1311

<210> 30
<211> 437
<212> PRT
<213> Caenorhabditis elegans

<400> 30
Lys Pro Lys Ala Ile Pro Ala Phe Asp Asn Ala Tyr Asp Ala Asp Phe
1               5                   10                  15

Asp Asn Ser Pro Pro Leu His His Tyr Ser Ala Val His Leu Glu Thr
            20                  25                  30

Gly Leu Ser Pro Leu Glu Glu Ala Gln Arg Ala Leu Arg Ala Asn Arg
        35                  40                  45
```

Ala Arg His Lys Pro Ser Asn Val Ser Leu Ala Glu Glu Ala Lys Leu
50 55 60

Ala Ala Arg Gln Arg Tyr Ser Asn Ala Ser Asp Ile Arg Arg Glu Glu
65 70 75 80

Glu Glu Glu Val Val Glu Glu Asp Pro Ala Val Val Val Pro Val Leu
85 90 95

Arg Lys Asp Leu Glu Val Glu Glu Ala Pro Lys Ser Val Arg Pro Pro
100 105 110

Arg Tyr Arg Lys Ser Arg Glu Ile Glu Glu Pro Val Val Val Asp Arg
115 120 125

Phe Val Glu Glu Glu Val Asp Glu Lys Glu Asp Ile Asp Ala Ile Phe
130 135 140

Glu Lys Tyr Arg Lys Thr Ser Val Ser Ala Asp Pro Lys Ser His Thr
145 150 155 160

Pro Ile Leu Met Ala Asp Glu Tyr Lys Glu Pro Gln Lys Gln Val Pro
165 170 175

Ala Pro Val Val Val Ala Gln Glu Ser Pro Ile Leu Lys Arg Arg Asn
180 185 190

Ser Leu Val Pro Ser Arg Ile Ser Gly Arg Gln Ser Thr Arg Arg Ser
195 200 205

Val Thr Ser Val Arg Ser Met Arg Gly Lys Arg Lys Thr Arg Ala Ile
210 215 220

Pro Glu Phe Phe Asp Leu Thr Arg His Gln Asn Ile Arg Leu Arg Ala
225 230 235 240

Pro Ala Thr Lys Lys Lys Arg Ile Ser Leu His Arg Val Glu Asp Thr
245 250 255

Glu Val Val Val Glu Leu Leu Asn Gly Gln Lys Val Glu Val Ala Cys
260 265 270

Arg Ser Asp Val Ile Ser Arg Asp Val Phe Ser Leu Ile Val Gln Asn
275 280 285

Met Asn Ile Asn Glu His Val Phe Phe Gly Leu Ser Phe Leu Arg Asp
290 295 300

Gly Glu His Tyr Phe Ile Glu Asp His Gln Arg Leu Glu Lys Phe Ala
305 310 315 320

Pro Ser Gly Trp Lys Ser Val Ala Arg Val Gly Val Lys Val Pro Tyr
325 330 335

Val Leu His Leu Arg Phe Lys Phe Tyr Pro Gln Ile Leu Asp Phe Ile
340 345 350

Lys Thr Asp Val Thr Met Asn Glu Leu Tyr Leu Gln Cys Arg Arg Asp

355 360 365

Val Leu Glu Glu Arg Ile Gln Pro Lys Arg Asp Ala Ala Phe Glu Leu
370 375 380

Ala Ala Leu Ala Leu Gln Ala Glu Phe Gly Asn Arg Pro Pro Pro Val
385 390 395 400

Ile Thr Asp Tyr Phe Asp Ile Gln His Tyr Leu Pro Lys Lys Tyr Ser
405 410 415

Ser Phe Glu Asp Gln Ser Arg Leu Lys Asn Ile Leu Ala Glu Leu His
420 425 430

Gly His Tyr Ala Gly
435

<210> 31
<211> 2574
<212> DNA
<213> Caenorhabditis elegans

<400> 31
atgcctgcta atgagctttt cggagattct gatccggaag gggatgaaat ttttgaaaga 60
gttcgaaagg tgcaaccaaa gtcgatcaat gtaactgaga atcaaaaggt agaccctatg 120
cgtaaagtca aaatcgagct tcaagctgtt ttggttgcgg aaaaaatacc aattagtaca 180
gaagaaatca gacggcgtct gctggattct tatggtgcat gtcctgatcc gaaaagatat 240
aactgttcga ctttggacga tcttcttcaa gcatgctcgg agtcaattgt acacactttc 300
ggccgtgatg gaatacaccg atatgggcca agaactactg aagccaacca agatattatc 360
gaaatggtcc agcagcagtc gagctcaaaa cgcccggccc gctcgttttt aggttcagga 420
gctactaata acctcagcac tcatgggtca tcattccggg cattcagagg tccgtatgcg 480
tcagaggaaa tcgctaaatc gagaggaaca cctgagcaat tcaaagcaag acacaagttg 540
ggtccagcaa aaacaatttc tcgcgtaaaa aaccttgcag aggttttgaa agaatatgct 600
gatgagatag gagtttcaca tcctgatgag ccaaatcgca agattgtaac actggcagct 660
cttgccaata agttcaaaca gttgtattgt ttaccagcat ggggaaagaa catatcggaa 720
agtgaactat acattcagct caatgttcct cctttcaacg aatatctgca tttctggcgt 780
cttagcgaaa aaggtgacat cttcgttgat tgtattgatc gtgacaatgc cgatccaact 840
caaaaaagtg aacaaaatcc gtcagcagat gtttctattc aatctgaatc ttttggcggt 900
aaaagttcag cttcagcgtt tgaacaatct gtagtatccg ctccttcaac tattagagat 960
caaacatccg attcctttga cgggttcaac agtttcgaag tgcccccaga aaatggaagc 1020
aaagattcaa aaattttcaa ctcgaatcaa gaaagcatcg atgactatcc aggaaatgct 1080
atatctcgag atcgaaccgc tgatatgacc gacattgcat tgcgctttgg aactgtctct 1140
gtggcaagcc aacaatgtcc ggtatcttcg tcactcgttc cacaaaatgg aattcttcgt 1200
cagtcgcgtg ctcaagaaga cgacaacaac acatctattc taactattca atcatctcgt 1260
cgcaatcatt cagtgcttcg tcatcgtacg atcaagcctc gcaatccaac acaaaatctt 1320
gctgaagtcg taaaaactca tggtagcatt ccttatgaag cgctttcgga ttgtgataag 1380
attatcgtcg acttaggaaa gaacattttc aaagtttatg caactcaacc tggagaaatg 1440
atggtccgcc tttgtgatcc ccacgttgac acgactacat tgccactcct agagaacaat 1500
cttcgggatc ctgtcgagtc tgatttacgt tggatgacac tgggaaattc ccatatcaag 1560
aaacaatctg ttaaagtggt caagcctgca atgtttattg cgccacgcgg atttttgttg 1620
attttaaaag atgaagaacg tgaagaaatg gacgttgaaa aggttgctac tatgggggaat 1680
atattgcggg ccgtgatggt agcgcctatt gtggagctcc agcgtgaaac cgtccgcacg 1740
ggatcagctg cagtctacgt ttatcgacaa ggagcggagg tgcgctatta tagagtgttg 1800
atagttggac aagctaaaca agatggtgaa gtgttggttt tgttggctga tgtcgatgat 1860
caatatttcg tggatgtaca tctatctcat ttgttcccga ttccggaaga agccagcttc 1920
aaacactttc cctcgaatgt tgtattcgca acattgcatg gagtactggg tttgacgctt 1980
tctgagcagg atgtgatgtt cgaaaacatt gataatgatg atacaaaacg atttgtgggt 2040

```
ggatattttc atggaaacga tgacagaatt cttaacatcg atatggtttg gaaaaatgag 2100
cgtggtcaat ttgagtggct ttcacaaatt gtcaagagac gtggagcagt tacttcgtcc 2160
gatgccaaca tcattcactt tccacattct gcactcgatg tgataaagtc tgttggacca 2220
gattgttcgg tgtgctttgt cgactattca gttcgtgatg aatctgcaac atcttcattg 2280
atggaatcga ctagaattgt tcatgatagt cgtgaatcta tgacaactac ttatgttggt 2340
gagatgccaa gcccaattat cgaagaaatc gatgcaacat cttcatttga cccaaaactc 2400
ttgaatctcc attcattgtt cgataagtta atcgaggaac agaatgttac tatgattgtg 2460
gggatgttcc aattcgttcg aagtttgaag gatttattcg gcgataacaa tgaatgggaa 2520
cgcttgctca cgtacatgct cacgactggt aaaaacaata acattcgtct ctga       2574
```

<210> 32
<211> 857
<212> PRT
<213> Caenorhabditis elegans

<400> 32

```
Met Pro Ala Asn Glu Leu Phe Gly Asp Ser Asp Pro Glu Gly Asp Glu
1               5                   10                  15

Ile Phe Glu Arg Val Arg Lys Val Gln Pro Lys Ser Ile Asn Val Thr
            20                  25                  30

Glu Asn Gln Lys Val Asp Pro Met Arg Lys Val Lys Ile Glu Leu Gln
        35                  40                  45

Ala Val Leu Val Ala Glu Lys Ile Pro Ile Ser Thr Glu Glu Ile Arg
    50                  55                  60

Arg Arg Leu Leu Asp Ser Tyr Gly Ala Cys Pro Asp Pro Lys Arg Tyr
65                  70                  75                  80

Asn Cys Ser Thr Leu Asp Asp Leu Leu Gln Ala Cys Ser Glu Ser Ile
                85                  90                  95

Val His Thr Phe Gly Arg Asp Gly Ile His Arg Tyr Gly Pro Arg Thr
            100                 105                 110

Thr Glu Ala Asn Gln Asp Ile Ile Glu Met Val Gln Gln Gln Ser Ser
        115                 120                 125

Ser Lys Arg Pro Ala Arg Ser Phe Leu Gly Ser Gly Ala Thr Asn Asn
    130                 135                 140

Leu Ser Thr His Gly Ser Ser Phe Arg Ala Phe Arg Gly Pro Tyr Ala
145                 150                 155                 160

Ser Glu Glu Ile Ala Lys Ser Arg Gly Thr Pro Glu Gln Phe Lys Ala
                165                 170                 175

Arg His Lys Leu Gly Pro Ala Lys Thr Ile Ser Arg Val Lys Asn Leu
            180                 185                 190

Ala Glu Val Leu Lys Glu Tyr Ala Asp Glu Ile Gly Val Ser His Pro
        195                 200                 205

Asp Glu Pro Asn Arg Lys Ile Val Thr Leu Ala Ala Leu Ala Asn Lys
    210                 215                 220
```

Phe Lys Gln Leu Tyr Cys Leu Pro Ala Trp Gly Lys Asn Ile Ser Glu
225 230 235 240

Ser Glu Leu Tyr Ile Gln Leu Asn Val Pro Pro Phe Asn Glu Tyr Leu
245 250 255

His Phe Trp Arg Leu Ser Glu Lys Gly Asp Ile Phe Val Asp Cys Ile
260 265 270

Asp Arg Asp Asn Ala Asp Pro Thr Gln Lys Ser Glu Gln Asn Pro Ser
275 280 285

Ala Asp Val Ser Ile Gln Ser Glu Ser Phe Gly Gly Lys Ser Ser Ala
290 295 300

Ser Ala Phe Glu Gln Ser Val Val Ser Ala Pro Ser Thr Ile Arg Asp
305 310 315 320

Gln Thr Ser Asp Ser Phe Asp Gly Phe Asn Ser Phe Glu Val Pro Pro
325 330 335

Glu Asn Gly Ser Lys Asp Ser Lys Ile Phe Asn Ser Asn Gln Glu Ser
340 345 350

Ile Asp Asp Tyr Pro Gly Asn Ala Ile Ser Arg Asp Arg Thr Ala Asp
355 360 365

Met Thr Asp Ile Ala Leu Arg Phe Gly Thr Val Ser Val Ala Ser Gln
370 375 380

Gln Cys Pro Val Ser Ser Ser Leu Val Pro Gln Asn Gly Ile Leu Arg
385 390 395 400

Gln Ser Arg Ala Gln Glu Asp Asp Asn Asn Thr Ser Ile Leu Thr Ile
405 410 415

Gln Ser Ser Arg Arg Asn His Ser Val Leu Arg His Arg Thr Ile Lys
420 425 430

Pro Arg Asn Pro Thr Gln Asn Leu Ala Glu Val Val Lys Thr His Gly
435 440 445

Ser Ile Pro Tyr Glu Ala Leu Ser Asp Cys Asp Lys Ile Ile Val Asp
450 455 460

Leu Gly Lys Asn Ile Phe Lys Val Tyr Ala Thr Gln Pro Gly Glu Met
465 470 475 480

Met Val Arg Leu Cys Asp Pro His Val Asp Thr Thr Thr Leu Pro Leu
485 490 495

Leu Glu Asn Asn Leu Arg Asp Pro Val Glu Ser Asp Leu Arg Trp Met
500 505 510

Thr Leu Gly Asn Ser His Ile Lys Lys Gln Ser Val Lys Val Val Lys
515 520 525

Pro Ala Met Phe Ile Ala Pro Arg Gly Phe Leu Leu Ile Leu Lys Asp

530 535 540

Glu Glu Arg Glu Glu Met Asp Val Glu Lys Val Ala Thr Met Gly Asn
545 550 555 560

Ile Leu Arg Ala Val Met Val Ala Pro Ile Val Glu Leu Gln Arg Glu
565 570 575

Thr Val Arg Thr Gly Ser Ala Ala Val Tyr Val Tyr Arg Gln Gly Ala
580 585 590

Glu Val Arg Tyr Tyr Arg Val Leu Ile Val Gly Gln Ala Lys Gln Asp
595 600 605

Gly Glu Val Leu Val Leu Leu Ala Asp Val Asp Asp Gln Tyr Phe Val
610 615 620

Asp Val His Leu Ser His Leu Phe Pro Ile Pro Glu Glu Ala Ser Phe
625 630 635 640

Lys His Phe Pro Ser Asn Val Val Phe Ala Thr Leu His Gly Val Leu
645 650 655

Gly Leu Thr Leu Ser Glu Gln Asp Val Met Phe Glu Asn Ile Asp Asn
660 665 670

Asp Asp Thr Lys Arg Phe Val Gly Gly Tyr Phe His Gly Asn Asp Asp
675 680 685

Arg Ile Leu Asn Ile Asp Met Val Trp Lys Asn Glu Arg Gly Gln Phe
690 695 700

Glu Trp Leu Ser Gln Ile Val Lys Arg Arg Gly Ala Val Thr Ser Ser
705 710 715 720

Asp Ala Asn Ile Ile His Phe Pro His Ser Ala Leu Asp Val Ile Lys
725 730 735

Ser Val Gly Pro Asp Cys Ser Val Cys Phe Val Asp Tyr Ser Val Arg
740 745 750

Asp Glu Ser Ala Thr Ser Ser Leu Met Glu Ser Thr Arg Ile Val His
755 760 765

Asp Ser Arg Glu Ser Met Thr Thr Thr Tyr Val Gly Glu Met Pro Ser
770 775 780

Pro Ile Ile Glu Glu Ile Asp Ala Thr Ser Ser Phe Asp Pro Lys Leu
785 790 795 800

Leu Asn Leu His Ser Leu Phe Asp Lys Leu Ile Glu Glu Gln Asn Val
805 810 815

Thr Met Ile Val Gly Met Phe Gln Phe Val Arg Ser Leu Lys Asp Leu
820 825 830

Phe Gly Asp Asn Asn Glu Trp Glu Arg Leu Leu Thr Tyr Met Leu Thr
835 840 845

```
Thr Gly Lys Asn Asn Asn Ile Arg Leu
    850                 855


<210> 33
<211> 1587
<212> DNA
<213> Caenorhabditis elegans

<400> 33
[illegible] gagttcgaaa ggtgcaacca aagtcgatca atgtaactga gaatcaaaag 60
[illegible] tgcgtaaagt caaaatcgag cttcaagctg ttttggttgc ggaaaaaata 120
[illegible] cagaagaaat cagacggcgt ctgctggatt cttatggtgc atgtcctgat 180
[illegible] ataactgttc gactttggac gatcttcttc aagcatgctc ggagtcaatt 240
[illegible] tcggccgtga tggaatacac cgatatgggc caagaactac tgaagccaac 300
[illegible] tcgaaatggt ccagcagcag tcgagctcaa aacgcccggc ccgctcgttt 360
[illegible] gagctactaa taacctcagc actcatgggt catcattccg ggcattcaga 420
[illegible] cgtcagagga aatcgctaaa tcgagaggaa cacctgagca attcaaagca 480
[illegible] tgggtccagc aaaaacaatt tctcgcgtaa aaaaccttgc agaggttttg 540
[illegible] ctgatgagat aggagtttca catcctgatg agccaaatcg caagattgta 600
[illegible] ctcttgccaa taagttcaaa cagttgtatt gtttaccagc atggggaaag 660
[illegible] aaagtgaact atacattcag ctcaatgttc ctcctttcaa cgaatatctg 720
catttctggc gtcttagcga aaaaggtgac atcttcgttg attgtattga tcgtgacaat 780
gccgatccaa ctcaaaaaag tgaacaaaat ccgtcagcag atgtttctat tcaatctgaa 840
tcttttggcg gtaaaagttc agcttcagcg tttgaacaat ctgtagtatc cgctccttca 900
actattagag atcaaacatc cgattccttt gacgggttca acagtttcga agtgccccca 960
gaaaatggaa gcaaagattc aaaaattttc aactcgaatc aagaaagcat cgatgactat 1020
ccaggaaatg ctatatctcg agatcgaacc gctgatatga ccgacattgc attgcgcttt 1080
ggaactgtct ctgtggcaag ccaacaatgt ccggtatctt cgtcactcgt tccacaaaat 1140
ggaattcttc gtcagtcgcg tgctcaagaa gacgacaaca acacatctat tctaactatt 1200
caatcatctc gtcgcaatca ttcagtgctt cgtcatcgta cgatcaagcc tcgcaatcca 1260
acacaaaatc ttgctgaagt cgtaaaaact catggtagca ttccttatga agcgctttcg 1320
gattgtgata agattatcgt cgacttagga aagaacattt tcaaagttta tgcaactcaa 1380
cctggagaaa tgatggtccg cctttgtgat ccccacgttg acacgactac attgccactc 1440
ctagagaaca atcttcggga tcctgtcgag tctgatttac gttggatgac actgggaaat 1500
tcccatatca agaaacaatc tgttaaagtg gtcaagcctg caatgtttat tgcgccacgc 1560
ggattttgt tgattttaaa agatgaa                                      1587

<210> 34
<211> 529
<212> PRT
<213> Caenorhabditis elegans

<400> 34
Ile Phe Glu Arg Val Arg Lys Val Gln Pro Lys Ser Ile Asn Val Thr
1               5                   10                  15

Glu Asn Gln Lys Val Asp Pro Met Arg Lys Val Lys Ile Glu Leu Gln
            20                  25                  30

Ala Val Leu Val Ala Glu Lys Ile Pro Ile Ser Thr Glu Glu Ile Arg
        35                  40                  45

Arg Arg Leu Leu Asp Ser Tyr Gly Ala Cys Pro Asp Pro Lys Arg Tyr
    50                  55                  60

Asn Cys Ser Thr Leu Asp Asp Leu Leu Gln Ala Cys Ser Glu Ser Ile
```

55

65 70 75 80

Val His Thr Phe Gly Arg Asp Gly Ile His Arg Tyr Gly Pro Arg Thr
85 90 95

Thr Glu Ala Asn Gln Asp Ile Ile Glu Met Val Gln Gln Gln Ser Ser
100 105 110

Ser Lys Arg Pro Ala Arg Ser Phe Leu Gly Ser Gly Ala Thr Asn Asn
115 120 125

Leu Ser Thr His Gly Ser Ser Phe Arg Ala Phe Arg Gly Pro Tyr Ala
130 135 140

Ser Glu Glu Ile Ala Lys Ser Arg Gly Thr Pro Glu Gln Phe Lys Ala
145 150 155 160

Arg His Lys Leu Gly Pro Ala Lys Thr Ile Ser Arg Val Lys Asn Leu
165 170 175

Ala Glu Val Leu Lys Glu Tyr Ala Asp Glu Ile Gly Val Ser His Pro
180 185 190

Asp Glu Pro Asn Arg Lys Ile Val Thr Leu Ala Ala Leu Ala Asn Lys
195 200 205

Phe Lys Gln Leu Tyr Cys Leu Pro Ala Trp Gly Lys Asn Ile Ser Glu
210 215 220

Ser Glu Leu Tyr Ile Gln Leu Asn Val Pro Pro Phe Asn Glu Tyr Leu
225 230 235 240

His Phe Trp Arg Leu Ser Glu Lys Gly Asp Ile Phe Val Asp Cys Ile
245 250 255

Asp Arg Asp Asn Ala Asp Pro Thr Gln Lys Ser Glu Gln Asn Pro Ser
260 265 270

Ala Asp Val Ser Ile Gln Ser Glu Ser Phe Gly Gly Lys Ser Ser Ala
275 280 285

Ser Ala Phe Glu Gln Ser Val Val Ser Ala Pro Ser Thr Ile Arg Asp
290 295 300

Gln Thr Ser Asp Ser Phe Asp Gly Phe Asn Ser Phe Glu Val Pro Pro
305 310 315 320

Glu Asn Gly Ser Lys Asp Ser Lys Ile Phe Asn Ser Asn Gln Glu Ser
325 330 335

Ile Asp Asp Tyr Pro Gly Asn Ala Ile Ser Arg Asp Arg Thr Ala Asp
340 345 350

Met Thr Asp Ile Ala Leu Arg Phe Gly Thr Val Ser Val Ala Ser Gln
355 360 365

Gln Cys Pro Val Ser Ser Ser Leu Val Pro Gln Asn Gly Ile Leu Arg
370 375 380

Gln Ser Arg Ala Gln Glu Asp Asp Asn Asn Thr Ser Ile Leu Thr Ile
385 390 395 400

Gln Ser Ser Arg Arg Asn His Ser Val Leu Arg His Arg Thr Ile Lys
405 410 415

Pro Arg Asn Pro Thr Gln Asn Leu Ala Glu Val Val Lys Thr His Gly
420 425 430

Ser Ile Pro Tyr Glu Ala Leu Ser Asp Cys Asp Lys Ile Ile Val Asp
435 440 445

Leu Gly Lys Asn Ile Phe Lys Val Tyr Ala Thr Gln Pro Gly Glu Met
450 455 460

Met Val Arg Leu Cys Asp Pro His Val Asp Thr Thr Thr Leu Pro Leu
465 470 475 480

Leu Glu Asn Asn Leu Arg Asp Pro Val Glu Ser Asp Leu Arg Trp Met
485 490 495

Thr Leu Gly Asn Ser His Ile Lys Lys Gln Ser Val Lys Val Val Lys
500 505 510

Pro Ala Met Phe Ile Ala Pro Arg Gly Phe Leu Leu Ile Leu Lys Asp
515 520 525

Glu

<210> 35
<211> 1593
<212> DNA
<213> Caenorhabditis elegans

<400> 35
atgcgcattg tccgtacaca ccgcgacgag tttttacgga cattgtgctt gaacttgttt 60
tgttgtctgt tgatcaactc aattgagaaa tcaaaacaaa ttcaatcttc agcgtatttt 120
tttcgaaatt ctcactcttt cgccatcgaa aagttcaaaa gaaaacaaca aaaaatgcct 180
cgcggtctac ggagagcaga tttagtcaaa cgacatcgcc actcaacgac aggagacaaa 240
gacggaggag taccagaagt aataggatgc ccagttttag atcctattat ctgccaatgt 300
ccaaaagatg agatcgagct tggtgaagga gtcaagatga cgtgcacttg ggaatcatgc 360
ccgtactcta gtagaccact tcatcacata tgctatcaac tgctcgagga caatcttgtc 420
aagcgattag cctcactggg aagtgcacga ggatggacag tgccacaacg gaggaataac 480
ttatgggaga ggaagggtca gtccctgatc ggaaagttct gccgatgtcg ctgcgatcgg 540
ggacaaatga ccagagacaa gcaggcttta tatgagaaag agaaggctgt ggaaaaagag 600
aagaagaaga aggccaagaa agcaaaacaa ctgccccagc tacaatttaa ttctaaacct 660
ttggcagcta tcgaggagaa aaagcgagga gacgctgatg tattccactc accgtccatt 720
gcctcaagta cacggcatca cacattctcg acgacgacac gatcgcgact tcatactgat 780
cgttcggctt cttccatttt aacacacact attggaagaa cgtggtccga atcttcgttt 840
gccggtgaaa caaatggtca gtacgacaac aatcaggagc cacatccatc aaattgtgaa 900
tgcgtatttc atcacgatta cgacgctgac gatcaaatag atacggattt cgagtgtgaa 960
agcaatcaca gcgacgtaat agttccagct ccacttccac cacttcaggc gaaaagctat 1020
gcagcgacaa taatgagaaa cgggacaccg aaggttacaa attattcacc ggatagtggt 1080
ctcgatcagc aaactccaag gttttcattg tcttcttcga gtggaggaga tgtcgataat 1140
caacatggag acttccacgt ggaaactaga atttccgagc atctcaacgc gttgggactc 1200

```
agcataatgt cgccggtgga gaatgcgaat gaaaatgtca attatgaaga atcgccgttc 1260
tacccggagc tgacatcgac tccaatcgtc tcgaagaagc agcgggaacc tctccgagcg 1320
aaaaagagca catctgtctc gaagcttcca cttgctccgt cgtcacagct attcaatgaa 1380
gaatcgcgtt gtggattcag attcaatgtg ccggttcgcg aaatgatgga catatggcaa 1440
gagtctggag ccttgtcgcc ggcaattcga gaaacacagg ctgaaaatac tgaaaaaaga 1500
gctgacaatg cgtcgggtgt actccaatat ggatggactc cattcttcgg caatggcttc 1560
aatctcggag agcgcctcta ctacttccca tag                             1593
```

<210> 36
<211> 330
<212> PRT
<213> Caenorhabditis elegans

<400> 36

```
Met Arg Ile Val Arg Thr His Arg Asp Glu Phe Leu Arg Thr Leu Cys
1               5                   10                  15

Leu Asn Leu Phe Cys Cys Leu Leu Ile Asn Ser Ile Glu Lys Ser Lys
            20                  25                  30

Asn Ile Gln Ser Ser Ala Tyr Phe Phe Arg Asn Ser His Ser Phe Ala
        35                  40                  45

Ile Glu Lys Phe Lys Arg Lys Gln Gln Lys Met Pro Arg Gly Leu Arg
    50                  55                  60

Arg Ala Asp Leu Val Lys Arg His Arg His Ser Thr Thr Gly Asp Lys
65                  70                  75                  80

Asp Gly Gly Val Pro Glu Val Ile Gly Cys Pro Val Leu Asp Pro Ile
                85                  90                  95

Ile Cys Gln Cys Pro Lys Asp Glu Ile Glu Leu Gly Glu Gly Val Lys
            100                 105                 110

Met Thr Cys Thr Trp Glu Ser Cys Pro Tyr Ser Ser Arg Pro Leu His
        115                 120                 125

His Ile Cys Tyr Gln Leu Leu Glu Asp Asn Leu Val Lys Arg Leu Ala
    130                 135                 140

Ser Leu Gly Ser Ala Arg Gly Trp Thr Val Pro Gln Arg Arg Asn Asn
145                 150                 155                 160

Leu Trp Glu Arg Lys Gly Gln Ser Leu Ile Gly Lys Phe Cys Arg Cys
                165                 170                 175

Arg Cys Asp Arg Gly Gln Met Thr Arg Asp Lys Gln Ala Leu Tyr Glu
            180                 185                 190

Lys Glu Lys Ala Val Glu Lys Glu Lys Lys Lys Lys Ala Lys Lys Ala
        195                 200                 205

Lys Gln Leu Pro Gln Leu Gln Phe Asn Ser Lys Pro Leu Ala Ala Ile
    210                 215                 220

Glu Glu Lys Lys Arg Gly Asp Ala Asp Val Phe His Ser Pro Ser Ile
225                 230                 235                 240
```

```
Ala Ser Ser Thr Arg His His Thr Phe Ser Thr Thr Thr Arg Ser Arg
                245                 250                 255

Leu His Thr Asp Arg Ser Ala Ser Ser Ile Leu Thr His Thr Ile Gly
            260                 265                 270

Arg Thr Trp Ser Glu Ser Ser Phe Ala Gly Glu Thr Asn Gly Gln Tyr
        275                 280                 285

Asp Asn Asn Gln Glu Pro His Pro Ser Asn Cys Glu Cys Val Phe His
    290                 295                 300

His Asp Tyr Asp Ala Asp Asp Gln Ile Asp Thr Asp Phe Glu Cys Glu
305                 310                 315                 320

Ser Asn His Ser Asp Val Ile Val Pro Ala Pro Leu Pro Pro Leu Gln
                325                 330                 335

Ala Lys Ser Tyr Ala Ala Thr Ile Met Arg Asn Gly Thr Pro Lys Val
            340                 345                 350

Thr Asn Tyr Ser Pro Asp Ser Gly Leu Asp Gln Gln Thr Pro Arg Phe
        355                 360                 365

Ser Leu Ser Ser Ser Ser Gly Gly Asp Val Asp Asn Gln His Gly Asp
    370                 375                 380

Phe His Val Glu Thr Arg Ile Ser Glu His Leu Asn Ala Leu Gly Leu
385                 390                 395                 400

Ser Ile Met Ser Pro Val Glu Asn Ala Asn Glu Asn Val Asn Tyr Glu
                405                 410                 415

Glu Ser Pro Phe Tyr Pro Glu Leu Thr Ser Thr Pro Ile Val Ser Lys
            420                 425                 430

Lys Gln Arg Glu Pro Leu Arg Ala Lys Lys Ser Thr Ser Val Ser Lys
        435                 440                 445

Leu Pro Leu Ala Pro Ser Ser Gln Leu Phe Asn Glu Glu Ser Arg Cys
    450                 455                 460

Gly Phe Arg Phe Asn Val Pro Val Arg Glu Met Met Asp Ile Trp Gln
465                 470                 475                 480

Glu Ser Gly Ala Leu Ser Pro Ala Ile Arg Glu Thr Gln Ala Glu Asn
                485                 490                 495

Thr Glu Lys Arg Ala Glu Asn Ala Ser Gly Val Leu Gln Tyr Gly Trp
            500                 505                 510

Thr Pro Phe Phe Gly Asn Gly Phe Asn Leu Gly Glu Arg Leu Tyr Tyr
        515                 520                 525

Phe Pro
    530
```

<210> 37
<211> 1458
<212> DNA
<213> Caenorhabditis elegans

<400> 37
tctttcgcca tcgaaaagtt caaaagaaaa caacaaaaaa tgcctcgcgg tctacggaga 60
gcagatttag tcaaacgaca tcgccactca acgacaggag acaaagacgg aggagtacca 120
gaagtaatag gatgcccagt tttagatcct attatctgcc aatgtccaaa agatgagatc 180
gagcttggtg aaggagtcaa gatgacgtgc acttgggaat catgcccgta ctctagtaga 240
ccacttcatc acatatgcta tcaactgctc gaggacaatc ttgtcaagcg attagcctca 300
ctgggaagtg cacgaggatg gacagtgcca caacggagga ataacttatg ggagaggaag 360
ggtcagtccc tgatcggaaa gttctgccga tgtcgctgcg atcggggaca aatgaccaga 420
gacaagcagg ctttatatga gaaagagaag gctgtggaaa aagagaagaa gaagaaggcc 480
aagaaagcaa aacaactgcc ccagctacaa tttaattcta aacctttggc agctatcgag 540
gagaaaaagc gaggagacgc tgatgtattc cactcaccgt ccattgcctc aagtacacgg 600
catcacacat tctcgacgac gacacgatcg cgacttcata ctgatcgttc ggcttcttcc 660
attttaacac acactattgg aagaacgtgg tccgaatctt cgtttgccgg tgaaacaaat 720
ggtcagtacg acaacaatca ggagccacat ccatcaaatt gtgaatgcgt atttcatcac 780
gattacgacg ctgacgatca aatagatacg gatttcgagt gtgaaagcaa tcacagcgac 840
gtaatagttc cagctccact tccaccactt caggcgaaaa gctatgcagc gacaataatg 900
agaaacggga caccgaaggt tacaaattat tcaccggata gtggtctcga tcagcaaact 960
ccaaggtttt cattgtcttc ttcgagtgga ggagatgtcg ataatcaaca tggagacttc 1020
cacgtggaaa ctagaatttc cgagcatctc aacgcgttgg gactcagcat aatgtcgccg 1080
gtggagaatg cgaatgaaaa tgtcaattat gaagaatcgc cgttctaccc ggagctgaca 1140
tcgactccaa tcgtctcgaa gaagcagcgg gaacctctcc gagcgaaaaa gagcacatct 1200
gtctcgaagc ttccacttgc tccgtcgtca cagctattca atgaagaatc gcgttgtgga 1260
ttcagattca atgtgccggt tcgcgaaatg atggacatat ggcaagagtc tggagccttg 1320
tcgccggcaa ttcgagaaac acaggctgaa aatactgaaa aaagagctga gaatgcgtcg 1380
ggtgtactcc aatatggatg gactccattc ttcggcaatg gcttcaatct cggagagcgc 1440
ctctactact tcccatag 1458

<210> 38
<211> 485
<212> PRT
<213> Caenorhabditis elegans

<400> 38
Ser Phe Ala Ile Glu Lys Phe Lys Arg Lys Gln Gln Lys Met Pro Arg
1               5                   10                  15

Gly Leu Arg Arg Ala Asp Leu Val Lys Arg His Arg His Ser Thr Thr
            20                  25                  30

Gly Asp Lys Asp Gly Gly Val Pro Glu Val Ile Gly Cys Pro Val Leu
        35                  40                  45

Asp Pro Ile Ile Cys Gln Cys Pro Lys Asp Glu Ile Glu Leu Gly Glu
    50                  55                  60

Gly Val Lys Met Thr Cys Thr Trp Glu Ser Cys Pro Tyr Ser Ser Arg
65                  70                  75                  80

Pro Leu His His Ile Cys Tyr Gln Leu Leu Glu Asp Asn Leu Val Lys
                85                  90                  95

Arg Leu Ala Ser Leu Gly Ser Ala Arg Gly Trp Thr Val Pro Gln Arg

60

100 105 110

Arg Asn Asn Leu Trp Glu Arg Lys Gly Gln Ser Leu Ile Gly Lys Phe
115 120 125

Cys Arg Cys Arg Cys Asp Arg Gly Gln Met Thr Arg Asp Lys Gln Ala
130 135 140

Leu Tyr Glu Lys Glu Lys Ala Val Glu Lys Glu Lys Lys Lys Lys Ala
145 150 155 160

Lys Lys Ala Lys Gln Leu Pro Gln Leu Gln Phe Asn Ser Lys Pro Leu
165 170 175

Ala Ala Ile Glu Glu Lys Lys Arg Gly Asp Ala Asp Val Phe His Ser
180 185 190

Pro Ser Ile Ala Ser Ser Thr Arg His His Thr Phe Ser Thr Thr Thr
195 200 205

Arg Ser Arg Leu His Thr Asp Arg Ser Ala Ser Ser Ile Leu Thr His
210 215 220

Thr Ile Gly Arg Thr Trp Ser Glu Ser Ser Phe Ala Gly Glu Thr Asn
225 230 235 240

Gly Gln Tyr Asp Asn Asn Gln Glu Pro His Pro Ser Asn Cys Glu Cys
245 250 255

Val Phe His His Asp Tyr Asp Ala Asp Asp Gln Ile Asp Thr Asp Phe
260 265 270

Glu Cys Glu Ser Asn His Ser Asp Val Ile Val Pro Ala Pro Leu Pro
275 280 285

Pro Leu Gln Ala Lys Ser Tyr Ala Ala Thr Ile Met Arg Asn Gly Thr
290 295 300

Pro Lys Val Thr Asn Tyr Ser Pro Asp Ser Gly Leu Asp Gln Gln Thr
305 310 315 320

Pro Arg Phe Ser Leu Ser Ser Ser Ser Gly Gly Asp Val Asp Asn Gln
325 330 335

His Gly Asp Phe His Val Glu Thr Arg Ile Ser Glu His Leu Asn Ala
340 345 350

Leu Gly Leu Ser Ile Met Ser Pro Val Glu Asn Ala Asn Glu Asn Val
355 360 365

Asn Tyr Glu Glu Ser Pro Phe Tyr Pro Glu Leu Thr Ser Thr Pro Ile
370 375 380

Val Ser Lys Lys Gln Arg Glu Pro Leu Arg Ala Lys Lys Ser Thr Ser
385 390 395 400

Val Ser Lys Leu Pro Leu Ala Pro Ser Ser Gln Leu Phe Asn Glu Glu
405 410 415

Ser Arg Cys Gly Phe Arg Phe Asn Val Pro Val Arg Glu Met Met Asp
420 425 430

Ile Trp Gln Glu Ser Gly Ala Leu Ser Pro Ala Ile Arg Glu Thr Gln
435 440 445

Ala Glu Asn Thr Glu Lys Arg Ala Glu Asn Ala Ser Gly Val Leu Gln
450 455 460

Tyr Gly Trp Thr Pro Phe Phe Gly Asn Gly Phe Asn Leu Gly Glu Arg
465 470 475 480

Leu Tyr Tyr Phe Pro
485

<210> 39
<211> 1056
<212> DNA
<213> Caenorhabditis elegans

<400> 39
atgcaaaaca cacagatatt tactaacttc gctcacagag cacatgatgg attaccgttt 60
aacagtgcga atccttccaa caaagatcca attttcacaa tgccaatatc ggtcaaaccg 120
aaaactcgtg agccggcttc tttcacagat aacaggatat atgtcagcaa tattcccttc 180
tcgtttcgtg aacaagattt ggcggcaatg ttcttcgcat atggaagagt cctgagtgtg 240
gaaatcgtca caaatgatcg tggatccaaa gggttcgggt ttgtcacact cgattccatc 300
gaatcctgtg agaaagctcg tgctgcgctt cacgaatcac atgttcaagg aagaattata 360
gaagtgagaa gagcgacacc aacccgcaga aagcttatca acaatccaca aaatgaagtt 420
ttgccaccac caaagctgtg tgtcgatctt cgagcccctc ataatttatg gagagctgag 480
ccaatgcatc agttgttcaa ggaaaaggag aacacaacat gttttcccga agctggattc 540
atgatggcac cataccgtag caatggaatt ttcaacacgc gtagtcttgt gcagaccaaa 600
ccacctcgat gcaccaagca cagcgagctc aagctttctt cagctggtga atacttctgc 660
aaaaacggcg agcctacgac ggaaacaagt attctgatgt gcatgcacag acaaaactca 720
ccatgcagca ataagtgttc tgattcttcg aatcacgagc tgtctgatgt ggagttgaac 780
tctatattcc cacatcatct tcgtgaccag attactgctc ttctcgacac ttcaaaccat 840
tttggatcag gaaataatag tgctaacaaa ggaaagagag caccatctgt gacatcttct 900
ggattgagat catcagagag cgagacagtt tcagacgaag agattcattg gtccccacat 960
aacagccctg attatcttct cgctgctctc tacgaaggtt ccacatcgtt ccacggaaag 1020
tctgtttctc caccaaaaga atcgtcaagc cagtaa 1056

<210> 40
<211> 351
<212> PRT
<213> Caenorhabditis elegans

<400> 40
Met Gln Asn Thr Gln Ile Phe Thr Asn Phe Ala His Arg Ala His Asp
1 5 10 15

Gly Leu Pro Phe Asn Ser Ala Asn Pro Ser Asn Lys Asp Pro Ile Phe
20 25 30

Thr Met Pro Ile Ser Val Lys Pro Lys Thr Arg Glu Pro Ala Ser Phe
35 40 45

Thr Asp Asn Arg Ile Tyr Val Ser Asn Ile Pro Phe Ser Phe Arg Glu

50 55 60

Gln Asp Leu Ala Ala Met Phe Phe Ala Tyr Gly Arg Val Leu Ser Val
65 70 75 80

Glu Ile Val Thr Asn Asp Arg Gly Ser Lys Gly Phe Gly Phe Val Thr
85 90 95

Leu Asp Ser Ile Glu Ser Cys Glu Lys Ala Arg Ala Ala Leu His Glu
100 105 110

Ser His Val Gln Gly Arg Ile Ile Glu Val Arg Arg Ala Thr Pro Thr
115 120 125

Arg Arg Lys Leu Ile Asn Asn Pro Gln Asn Glu Val Leu Pro Pro Pro
130 135 140

Lys Leu Cys Val Asp Leu Arg Ala Pro His Asn Leu Trp Arg Ala Glu
145 150 155 160

Pro Met His Gln Leu Phe Lys Glu Lys Glu Asn Thr Thr Cys Phe Pro
165 170 175

Glu Ala Gly Phe Met Met Ala Pro Tyr Arg Ser Asn Gly Ile Phe Asn
180 185 190

Thr Arg Ser Leu Val Gln Thr Lys Pro Pro Arg Cys Thr Lys His Ser
195 200 205

Glu Leu Lys Leu Ser Ser Ala Gly Glu Tyr Phe Cys Lys Asn Gly Glu
210 215 220

Pro Thr Thr Glu Thr Ser Ile Leu Met Cys Met His Arg Gln Asn Ser
225 230 235 240

Pro Cys Ser Asn Lys Cys Ser Asp Ser Ser Asn His Glu Leu Ser Asp
245 250 255

Val Glu Leu Asn Ser Ile Phe Pro His His Leu Arg Asp Gln Ile Thr
260 265 270

Ala Leu Leu Asp Thr Ser Asn His Phe Gly Ser Gly Asn Asn Ser Ala
275 280 285

Asn Lys Gly Lys Arg Ala Pro Ser Val Thr Ser Ser Gly Leu Arg Ser
290 295 300

Ser Glu Ser Glu Thr Val Ser Asp Glu Glu Ile His Trp Ser Pro His
305 310 315 320

Asn Ser Pro Asp Tyr Leu Leu Ala Ala Leu Tyr Glu Gly Ser Thr Ser
325 330 335

Phe His Gly Lys Ser Val Ser Pro Pro Lys Glu Ser Ser Ser Gln
340 345 350

<210> 41

<211> 1053
<212> DNA
<213> Caenorhabditis elegans

<400> 41
caaaacacac agatatttac taacttcgct cacagagcac atgatggatt accgtttaac 60
agtgcgaatc cttccaacaa agatccaatt ttcacaatgc caatatcggt caaaccgaaa 120
actcgtgagc cggcttcttt cacagataac aggatatatg tcagcaatat tcccttctcg 180
tttcgtgaac aagatttggc ggcaatgttc ttcgcatatg gaagagtcct gagtgtggaa 240
atcgtcacaa atgatcgtgg atccaaaggg ttcgggtttg tcacactcga ttccatcgaa 300
tcctgtgaga aagctcgtgc tgcgcttcac gaatcacatg ttcaaggaag aattatagaa 360
gtgagaagag cgacaccaac ccgcagaaag cttatcaaca atccacaaaa tgaagttttg 420
ccaccaccaa agctgtgtgt cgatcttcga gcccctcata atttatggag agctgagcca 480
atgcatcagt tgttcaagga aaaggagaac acaacatgtt ttcccgaagc tggattcatg 540
atggcaccat accgtagcaa tggaattttc aacacgcgta gtcttgtgca gaccaaacca 600
cctcgatgca ccaagcacag cgagctcaag ctttcttcag ctggtgaata cttctgcaaa 660
aacggcgagc ctacgacgga aacaagtatt ctgatgtgca tgcacagaca aaactcacca 720
tgcagcaata agtgttctga ttcttcgaat cacgagctgt ctgatgtgga gttgaactct 780
atattcccac atcatcttcg tgaccagatt actgctcttc tcgacacttc aaaccatttt 840
ggatcaggaa ataatagtgc taacaaagga aagagagcac catctgtgac atcttctgga 900
ttgagatcat cagagagcga gacagtttca gacgaagaga ttcattggtc cccacataac 960
agccctgatt atcttctcgc tgctctctac gaaggttcca catcgttcca cggaaagtct 1020
gtttctccac caaaagaatc gtcaagccag taa 1053

<210> 42
<211> 350
<212> PRT
<213> Caenorhabditis elegans

<400> 42
Gln Asn Thr Gln Ile Phe Thr Asn Phe Ala His Arg Ala His Asp Gly
1 5 10 15

Leu Pro Phe Asn Ser Ala Asn Pro Ser Asn Lys Asp Pro Ile Phe Thr
20 25 30

Met Pro Ile Ser Val Lys Pro Lys Thr Arg Glu Pro Ala Ser Phe Thr
35 40 45

Asp Asn Arg Ile Tyr Val Ser Asn Ile Pro Phe Ser Phe Arg Glu Gln
50 55 60

Asp Leu Ala Ala Met Phe Phe Ala Tyr Gly Arg Val Leu Ser Val Glu
65 70 75 80

Ile Val Thr Asn Asp Arg Gly Ser Lys Gly Phe Gly Phe Val Thr Leu
85 90 95

Asp Ser Ile Glu Ser Cys Glu Lys Ala Arg Ala Ala Leu His Glu Ser
100 105 110

His Val Gln Gly Arg Ile Ile Glu Val Arg Arg Ala Thr Pro Thr Arg
115 120 125

Arg Lys Leu Ile Asn Asn Pro Gln Asn Glu Val Leu Pro Pro Pro Lys
130 135 140

Leu Cys Val Asp Leu Arg Ala Pro His Asn Leu Trp Arg Ala Glu Pro

```
145                 150                 155                 160

Met His Gln Leu Phe Lys Glu Lys Glu Asn Thr Thr Cys Phe Pro Glu
                165                 170                 175

Ala Gly Phe Met Met Ala Pro Tyr Arg Ser Asn Gly Ile Phe Asn Thr
            180                 185                 190

Arg Ser Leu Val Gln Thr Lys Pro Pro Arg Cys Thr Lys His Ser Glu
        195                 200                 205

Leu Lys Leu Ser Ser Ala Gly Glu Tyr Phe Cys Lys Asn Gly Glu Pro
    210                 215                 220

Thr Thr Glu Thr Ser Ile Leu Met Cys Met His Arg Gln Asn Ser Pro
225                 230                 235                 240

Cys Ser Asn Lys Cys Ser Asp Ser Ser Asn His Glu Leu Ser Asp Val
                245                 250                 255

Glu Leu Asn Ser Ile Phe Pro His His Leu Arg Asp Gln Ile Thr Ala
            260                 265                 270

Leu Leu Asp Thr Ser Asn His Phe Gly Ser Gly Asn Asn Ser Ala Asn
        275                 280                 285

Lys Gly Lys Arg Ala Pro Ser Val Thr Ser Ser Gly Leu Arg Ser Ser
    290                 295                 300

Glu Ser Glu Thr Val Ser Asp Glu Glu Ile His Trp Ser Pro His Asn
305                 310                 315                 320

Ser Pro Asp Tyr Leu Leu Ala Ala Leu Tyr Glu Gly Ser Thr Ser Phe
                325                 330                 335

His Gly Lys Ser Val Ser Pro Pro Lys Glu Ser Ser Ser Gln
            340                 345                 350
```

<210> 43
<211> 1349
<212> DNA
<213> Caenorhabditis elegans

<400> 43

```
aagtttgagg taacaaaaat gccgccgcca gctccgttcg aggggcaaaa gaacaaaggg 60
cacaatgtcc acaaaactgg aggaaaagca anaaagcgaa atgagaagga gccgagagtc 120
aagggaaaca atctgaaagc tttcactttc cactccgccg tatccgctgg aaaagcgatt 180
cgacgagctg cagatctcaa tgaaaagaag aaacatgttc tgatgatgga cagaaaaccc 240
atcgaaacac caccaatcat tgtagcaatc gttggaccga gtaaagtcgg aaaaacgaca 300
cttctccggg gtcttgtcaa gtattacctc cgtgatggat tcggagagat caatggtcca 360
gtgacaattg taactggaaa gaaacgtcgt gtacagttca ttgaggtcaa aaacgatatt 420
aatcatatga ttgatatcgc gaaagtcgca gatttggtgc ttctaatggt cgatgcatcg 480
tatggatttg aaatggaaac ctttgaattt ctaaatattt gccaagtgca cggaatgcca 540
cgtattatgg gagtattgaa tcatttggat cttctcgatg gaatctcacg tgtcaataag 600
accaagaaaa ttctgaaaca tcgtttctgg acggagctct accagggcgc gaagcttttc 660
tacatgactg gaatgatgca tggacagtat aaatataatg agatccataa cctctgcaga 720
ttcatttctg tcatgaaatt ccgtccgatg gtgtggaaag atgctcatcc atacgttctt 780
```

```
tgtgatcgtt tcgaagacat taccaacgtc gaaactcttc gaacggatcc actcatcgat 840
cgacacattg ccatgtatgg atgggttcat ggtgctcatt tgaagaatca ttcgtcgatt 900
catgtgccag gtgttggtga tatgaggatc agtaatgtca cgagtctacc cgatccgtgt 960
ccgctgcctg atgagattaa gaaacgagcg ttgaatgaga aagagcggaa agtgtatgct 1020
ccgtttttctg gattaggagg tgtcatttat gataaggatg cgatttatat tgagtcaaag 1080
aatgctcaca attttaatag aaaacgcgac ggactcgttg aagctctcga aggcgtcaag 1140
tcaggaaccg atgataaatt gaagaaatcc tctctgcaac ttctcggtga ttcagtagca 1200
cttgatattg atcaggaaag tgattggcca gagcctggag aagaagatga agaagatctg 1260
gatgaggagg attttcagga tgaagaagaa gatgaagatg aggatgagga tgaggaagat 1320
gttggtgtcg tgaaaaagga aggtgtact                                   1349

<210> 44
<211> 449
<212> PRT
<213> Caenorhabditis elegans

<400> 44
Lys Phe Glu Val Thr Lys Met Pro Pro Pro Ala Pro Phe Glu Gly Gln
1               5                   10                  15

Lys Asn Lys Gly His Asn Val His Lys Thr Gly Gly Lys Ala Xaa Lys
            20                  25                  30

Arg Asn Glu Lys Glu Pro Arg Val Lys Gly Asn Asn Leu Lys Ala Phe
        35                  40                  45

Thr Phe His Ser Ala Val Ser Ala Gly Lys Ala Ile Arg Arg Ala Ala
    50                  55                  60

Asp Leu Asn Glu Lys Lys Lys His Val Leu Met Met Asp Arg Lys Pro
65                  70                  75                  80

Ile Glu Thr Pro Pro Ile Ile Val Ala Ile Val Gly Pro Ser Lys Val
                85                  90                  95

Gly Lys Thr Thr Leu Leu Arg Gly Leu Val Lys Tyr Tyr Leu Arg Asp
            100                 105                 110

Gly Phe Gly Glu Ile Asn Gly Pro Val Thr Ile Val Thr Gly Lys Lys
        115                 120                 125

Arg Arg Val Gln Phe Ile Glu Val Lys Asn Asp Ile Asn His Met Ile
    130                 135                 140

Asp Ile Ala Lys Val Ala Asp Leu Val Leu Leu Met Val Asp Ala Ser
145                 150                 155                 160

Tyr Gly Phe Glu Met Glu Thr Phe Glu Phe Leu Asn Ile Cys Gln Val
                165                 170                 175

His Gly Met Pro Arg Ile Met Gly Val Leu Asn His Leu Asp Leu Leu
            180                 185                 190

Asp Gly Ile Ser Arg Val Asn Lys Thr Lys Lys Ile Leu Lys His Arg
        195                 200                 205

Phe Trp Thr Glu Leu Tyr Gln Gly Ala Lys Leu Phe Tyr Met Thr Gly
    210                 215                 220
```

Met Met His Gly Gln Tyr Lys Tyr Asn Glu Ile His Asn Leu Cys Arg
225 230 235 240

Phe Ile Ser Val Met Lys Phe Arg Pro Met Val Trp Lys Asp Ala His
245 250 255

Pro Tyr Val Leu Cys Asp Arg Phe Glu Asp Ile Thr Asn Val Glu Thr
260 265 270

Leu Arg Thr Asp Pro Leu Ile Asp Arg His Ile Ala Met Tyr Gly Trp
275 280 285

Val His Gly Ala His Leu Lys Asn His Ser Ser Ile His Val Pro Gly
290 295 300

Val Gly Asp Met Arg Ile Ser Asn Val Thr Ser Leu Pro Asp Pro Cys
305 310 315 320

Ile Leu Pro Asp Glu Ile Lys Lys Arg Ala Leu Asn Glu Lys Glu Arg
325 330 335

Lys Val Tyr Ala Pro Phe Ser Gly Leu Gly Gly Val Ile Tyr Asp Lys
340 345 350

Asp Ala Ile Tyr Ile Glu Ser Lys Asn Ala His Asn Phe Asn Arg Lys
355 360 365

Arg Asp Gly Leu Val Glu Ala Leu Glu Gly Val Lys Ser Gly Thr Asp
370 375 380

Asp Lys Leu Lys Lys Ser Ser Leu Gln Leu Leu Gly Asp Ser Val Ala
385 390 395 400

Leu Asp Ile Asp Gln Glu Ser Asp Trp Pro Glu Pro Gly Glu Glu Asp
405 410 415

Glu Glu Asp Leu Asp Glu Glu Asp Phe Gln Asp Glu Glu Glu Asp Glu
420 425 430

Asp Glu Asp Glu Asp Glu Glu Asp Val Gly Val Val Lys Lys Glu Gly
435 440 445

Val

<210> 45
<211> 3423
<212> DNA
<213> Caenorhabditis elegans

<400> 45
atgtcgtacc gcgtgttttc acgcgtcgat ccccatgtgc catcgacatc cacagctcca 60
cagcgtcgtt ttcaagaaaa cttgcaaaaa ttcaagcgat tttttccgcc aaacgcacca 120
ccaacactct gggaagtttg ctcctcgaaa caacgcagca aatcattgaa aaacacgttt 180
caaacggaag tacgtgcact acgaggactt aattttacag tattgctgaa tccgtacaaa 240
aactatctca atgatctcac aaatctatcc ggtttcacct tcgatgatct ttgtcaagca 300

```
cttcgattct ttgcatttta tagaaaacag ccagttttga agtcaaatat ggaagatgct 360
aacgaattat ttcgattaat tgcaagttgc atcatttatt caaatgataa ctggagggcg 420
tccatcgata aatcaacact agtggatacg ctgtcaatga acattttgga gaagcagagg 480
cttaagaatc tgaaacaaga atcatctgaa caaaaagatc caatataccc accactattc 540
caagatgatg agctaccttc tgttccaata caaattggca gattgaaaga cagagaaaaa 600
gtaccaattc ctccgcctcc ttgtagaaat gattttcga tgcgacaatt taatccgttg 660
gaagatgaac atttacgatc aatgcacctg tggaatcatg ttgggtgtaa tgatgcaaaa 720
ttcaatgggc catttgagag aactatcaaa atgatgtcta agaataatgt tgcgattcgc 780
tccaaagatc gacgattgag cgatgtggaa tattacggag acaatgaaga tttaccatcc 840
acacatatca gttttagatt agattctgtg atgcaattga ttaactttga tttccccaaa 900
attgaagatg atggatattt ctccaaagaa tgtctggata gtgcatggta tctctacgag 960
aattatcaaa ctgctcttca tgaatgcaca actgctttcg cagttatccg acctccatct 1020
ggacgtacaa tcaagccggg attcgttgaa gatggcctaa ctacagacga atgttcagaa 1080
tttcatatga tgggacgaca tattcatgga tttttccaag tatggaggga agaggatcgt 1140
ggatggagag aatcaaatgg aaaatgggtc ccacggcgat atttggtgga tatctacaat 1200
catattatgt ttccattatt cgtaaaatgg gaactttggc cgtcaacatt gaaatgggct 1260
tttgataagt acagcttgta tggattaaga ctaatgtcaa tgattagaag acatccacaa 1320
gaattgttga atgctggaga gaacttattt tcgaggtatc catcgcattt gttggaatcg 1380
aatagatatg atatgagcac tacgaaagga cggaatcagt atttatcagc aattcaaatg 1440
gagaataatc gagttgtgga caagcatatg cacagttcgg cttacaagct tcttatcgaa 1500
gaggacggaa gaaggagaaa gaggaaaccg aaggatgaag ctctactcgg agttgctgca 1560
aaagtacgaa cacctcgaaa agttcttgag ccacctctat ttgctccaac tcgatttatt 1620
tcttcatcaa ctccaaaaca acgcgccctg cttgttcaaa aagagaattt ggagaaaaca 1680
atgattaatc aagtgccacc agttgtaaat acaccaccat ccccacaaca gactgcttca 1740
caactcaaaa agacgccaac ttcggctaca aaacgtcatt taccggaaat tgagcaagaa 1800
ttgaagtcgg aatctgtccc agctccacca ccaaccaaga aaatgtcaat tatagctgac 1860
agttgggatg atcatgttgg taattcgatg gaagaagaac atgtagatga gaaggattcg 1920
gagaaaatgg aagattcaga aggaagacag aatgtttggg ttccacagga tcgaggaaaa 1980
gaatatgcac ctgaacaata tgcgcgagat attatcgaac attatattcc tgctgctcga 2040
gatcatccac ctcaaccaca acaaccacca cctccactac caacgccgaa gcctccacga 2100
agacggaaat ccggtcagaa aactgatcaa acgactccat catcagacgc cgaagcttca 2160
tccgatcctg caccgcctgt tcctgctgct ccagtggctc ctgttgttcc aattgttcca 2220
attgttccag ttcatcctgt acctttgcca aacggaagtg taaatactcc aaaagtgaag 2280
acgattgcaa agacaacagc acgagtactg tattccatta aacctcaaat accaccaatt 2340
gcgaacaaaa ctgtgtatcc tgtcaagaag ttgacacctt ctgtagttcc gtctccaatg 2400
attttaaatg gaaatactgc aactgcaagt ccatcgaaaa atgcagcatc tgtagttgtc 2460
agaaatgcat atacattttc attacagcaa aaagctccat attatccagc tggaatgcgt 2520
ccgaaaccaa ctcaaaatgg aattgaaaca cctccgacag gagcacaatc tttaatgcgt 2580
gcagccttct atagtgagag tcatcctaca cgaagtcccc tcgtcccgta tggttttgtt 2640
cctccagttg ccacatcatc tacttttgta cctgctgcaa caataccatc acctgctagc 2700
cgagcaatag ctcatcagaa gcaaatgctc ttaaatacgg aaacgtgtcg acgcgttatg 2760
ccgtttaata ttcaaatggc attcaaacca cgtcgttggg atcctcttcc gaaatcttct 2820
ggtgttttgg ctcattcaaa ttcaacaata ccatatgttc agcgagtccc caacaactca 2880
acacaatcag actttcggcc gagaagcttc agtcaaaata gtgttgcttc cccggctcct 2940
gctcctgttc caaatgcaat aaagagaaga gaagtaggga atttaaaatc tcggcagtat 3000
gtgccatgga tcgcaaatag cagggcgttg gtggcagcgg cgatggcaac gatggaggag 3060
acggctgaaa agatgtcttc gagcccgtta ctatcgtcac aggctccgat gacaactttg 3120
atgccaactc ctcctcctcc ggctccggct cctgctcaag catctgcaca atcaacatca 3180
gcaactccag cattggtgga tacaatatcg gctggttcaa cgactacaga aacaactaca 3240
ggcgattcaa atcagtcaaa tcctccttta cggacgtaca catcacacat tcgaaaaact 3300
cctggaacaa cactgacacc tgaagaaatc ggcgacgcga ttcgaactga aagtcaaaga 3360
tttcaagagg atggtgatga agggccaaca gtgaaaagct tcctaatgaa catctacaaa 3420
tga 3423
```

<210> 46
<211> 1140
<212> PRT
<213> Caenorhabditis elegans

<400> 46

Met Ser Tyr Arg Val Phe Ser Arg Val Asp Pro His Val Pro Ser Thr
1 5 10 15

Ser Thr Ala Pro Gln Arg Arg Phe Gln Glu Asn Leu Gln Lys Phe Lys
20 25 30

Arg Phe Phe Pro Pro Asn Ala Pro Pro Thr Leu Trp Glu Val Cys Ser
35 40 45

Ser Lys Gln Arg Ser Lys Ser Leu Lys Asn Thr Phe Gln Thr Glu Val
50 55 60

Arg Ala Leu Arg Gly Leu Asn Phe Thr Val Leu Leu Asn Pro Tyr Lys
65 70 75 80

Asn Tyr Leu Asn Asp Leu Thr Asn Leu Ser Gly Phe Thr Phe Asp Asp
85 90 95

Leu Cys Gln Ala Leu Arg Phe Phe Ala Phe Tyr Arg Lys Gln Pro Val
100 105 110

Leu Lys Ser Asn Met Glu Asp Ala Asn Glu Leu Phe Arg Leu Ile Ala
115 120 125

Ser Cys Ile Ile Tyr Ser Asn Asp Asn Trp Arg Ala Ser Ile Asp Lys
130 135 140

Ser Thr Leu Val Asp Thr Leu Ser Met Asn Ile Leu Glu Lys Gln Arg
145 150 155 160

Leu Lys Asn Leu Lys Gln Glu Ser Ser Glu Gln Lys Asp Pro Ile Tyr
165 170 175

Pro Pro Leu Phe Gln Asp Asp Glu Leu Pro Ser Val Pro Ile Gln Ile
180 185 190

Gly Arg Leu Lys Asp Arg Glu Lys Val Pro Ile Pro Pro Pro Pro Cys
195 200 205

Arg Asn Asp Phe Ser Met Arg Gln Phe Asn Pro Leu Glu Asp Glu His
210 215 220

Leu Arg Ser Met His Leu Trp Asn His Val Gly Cys Asn Asp Ala Lys
225 230 235 240

Phe Asn Gly Pro Phe Glu Arg Thr Ile Lys Met Met Ser Lys Asn Asn
245 250 255

Val Ala Ile Arg Ser Lys Asp Arg Arg Leu Ser Asp Val Glu Tyr Tyr
260 265 270

Gly Asp Asn Glu Asp Leu Pro Ser Thr His Ile Ser Phe Arg Leu Asp
275 280 285

Ser Val Met Gln Leu Ile Asn Phe Asp Phe Pro Lys Ile Glu Asp Asp
290 295 300

Gly Tyr Phe Ser Lys Glu Cys Leu Asp Ser Ala Trp Tyr Leu Tyr Glu
305 310 315 320

Asn Tyr Gln Thr Ala Leu His Glu Cys Thr Thr Ala Phe Ala Val Ile
325 330 335

Arg Pro Pro Ser Gly Arg Thr Ile Lys Pro Gly Phe Val Glu Asp Gly
340 345 350

Leu Thr Thr Asp Glu Cys Ser Glu Phe His Met Met Gly Arg His Ile
355 360 365

His Gly Phe Phe Gln Val Trp Arg Glu Glu Asp Arg Gly Trp Arg Glu
370 375 380

[illegible] Asn Gly Lys Trp Val Pro Arg Arg Tyr Leu Val Asp Ile Tyr Asn
385 390 395 400

His Ile Met Phe Pro Leu Phe Val Lys Trp Glu Leu Trp Pro Ser Thr
405 410 415

Leu Lys Trp Ala Phe Asp Lys Tyr Ser Leu Tyr Gly Leu Arg Leu Met
420 425 430

Ser Met Ile Arg Arg His Pro Gln Glu Leu Leu Asn Ala Gly Glu Asn
435 440 445

Leu Phe Ser Arg Tyr Pro Ser His Leu Leu Glu Ser Asn Arg Tyr Asp
450 455 460

Met Ser Thr Thr Lys Gly Arg Asn Gln Tyr Leu Ser Ala Ile Gln Met
465 470 475 480

Glu Asn Asn Arg Val Val Asp Lys His Met His Ser Ser Ala Tyr Lys
485 490 495

Leu Leu Ile Glu Glu Asp Gly Arg Arg Arg Lys Arg Lys Pro Lys Asp
500 505 510

Glu Ala Leu Leu Gly Val Ala Ala Lys Val Arg Thr Pro Arg Lys Val
515 520 525

Leu Glu Pro Pro Leu Phe Ala Pro Thr Arg Phe Ile Ser Ser Ser Thr
530 535 540

Pro Lys Gln Arg Ala Leu Leu Val Gln Lys Glu Asn Leu Glu Lys Thr
545 550 555 560

Met Ile Asn Gln Val Pro Pro Val Val Asn Thr Pro Pro Ser Pro Gln
565 570 575

Gln Thr Ala Ser Gln Leu Lys Lys Thr Pro Thr Ser Ala Thr Lys Arg
580 585 590

His Leu Pro Glu Ile Glu Gln Glu Leu Lys Ser Glu Ser Val Pro Ala
595 600 605

Pro Pro Pro Thr Lys Lys Met Ser Ile Ile Ala Asp Ser Trp Asp Asp
610 615 620

His Val Gly Asn Ser Met Glu Glu Glu His Val Asp Glu Lys Asp Ser
625 630 635 640

Glu Lys Met Glu Asp Ser Glu Gly Arg Gln Asn Val Trp Val Pro Gln
645 650 655

Asp Arg Gly Lys Glu Tyr Ala Pro Glu Gln Tyr Ala Arg Asp Ile Ile
660 665 670

Glu His Tyr Ile Pro Ala Ala Arg Asp His Pro Pro Gln Pro Gln Gln
675 680 685

Pro Pro Pro Pro Leu Pro Thr Pro Lys Pro Pro Arg Arg Arg Lys Ser
690 695 700

Gly Gln Lys Thr Asp Gln Thr Thr Pro Ser Ser Asp Ala Glu Ala Ser
705 710 715 720

Ser Asp Pro Ala Pro Pro Val Pro Ala Ala Pro Val Ala Pro Val Val
725 730 735

Pro Ile Val Pro Ile Val Pro Val His Pro Val Pro Leu Pro Asn Gly
740 745 750

Ser Val Asn Thr Pro Lys Val Lys Thr Ile Ala Lys Thr Thr Ala Arg
755 760 765

Val Leu Tyr Ser Ile Lys Pro Gln Ile Pro Pro Ile Ala Asn Lys Thr
770 775 780

Val Tyr Pro Val Lys Lys Leu Thr Pro Ser Val Val Pro Ser Pro Met
785 790 795 800

Ile Leu Asn Gly Asn Thr Ala Thr Ala Ser Pro Ser Lys Asn Ala Ala
805 810 815

Ser Val Val Val Arg Asn Ala Tyr Thr Phe Ser Leu Gln Gln Lys Ala
820 825 830

Pro Tyr Tyr Pro Ala Gly Met Arg Pro Lys Pro Thr Gln Asn Gly Ile
835 840 845

Glu Thr Pro Pro Thr Gly Ala Gln Ser Leu Met Arg Ala Ala Phe Tyr
850 855 860

Ser Glu Ser His Pro Thr Arg Ser Pro Leu Val Pro Tyr Gly Phe Val
865 870 875 880

Pro Pro Val Ala Thr Ser Ser Thr Phe Val Pro Ala Ala Thr Ile Pro
885 890 895

Ser Pro Ala Ser Arg Ala Ile Ala His Gln Lys Gln Met Leu Leu Asn
900 905 910

Thr Glu Thr Cys Arg Arg Val Met Pro Phe Asn Ile Gln Met Ala Phe

915 920 925

Lys Pro Arg Arg Trp Asp Pro Leu Pro Lys Ser Ser Gly Val Leu Ala
930 935 940

His Ser Asn Ser Thr Ile Pro Tyr Val Gln Arg Val Pro Asn Asn Ser
945 950 955 960

Thr Gln Ser Asp Phe Arg Pro Arg Ser Phe Ser Gln Asn Ser Val Ala
965 970 975

Ser Pro Ala Pro Ala Pro Val Pro Asn Ala Ile Lys Arg Arg Glu Val
980 985 990

Gly Asn Leu Lys Ser Arg Gln Tyr Val Pro Trp Ile Ala Asn Ser Arg
995 1000 1005

Ala Leu Val Ala Ala Ala Met Ala Thr Met Glu Glu Thr Ala Glu Lys
1010 1015 1020

Met Ser Ser Ser Pro Leu Leu Ser Ser Gln Ala Pro Met Thr Thr Leu
1025 1030 1035 1040

Met Pro Thr Pro Pro Pro Pro Ala Pro Ala Pro Ala Gln Ala Ser Ala
1045 1050 1055

Gln Ser Thr Ser Ala Thr Pro Ala Leu Val Asp Thr Ile Ser Ala Gly
1060 1065 1070

Ser Thr Thr Thr Glu Thr Thr Thr Gly Asp Ser Asn Gln Ser Asn Pro
1075 1080 1085

Pro Leu Arg Thr Tyr Thr Ser His Ile Arg Lys Thr Pro Gly Thr Thr
1090 1095 1100

Leu Thr Pro Glu Glu Ile Gly Asp Ala Ile Arg Thr Glu Ser Gln Arg
1105 1110 1115 1120

Phe Gln Glu Asp Gly Asp Glu Gly Pro Thr Val Lys Ser Phe Leu Met
1125 1130 1135

Asn Ile Tyr Lys
1140

<210> 47
<211> 1644
<212> DNA
<213> Caenorhabditis elegans

<400> 47
ttaccggaaa ttgagcaaga attgaagtcg gaatctgtcc cagctccacc accaaccaag 60
aaaatgtcaa ttatagctga cagttgggat gatcatgttg gtaattcgat ggaagaagaa 120
catgtagatg agaaggattc ggagaaaatg gaagattcag aaggaagaca gaatgtttgg 180
gttccacagg atcgaggaaa agaatatgca cctgaacaat atgcgcgaga tattatcgaa 240
cattatattc ctgctgctcg agatcatcca cctcaaccac aacaaccacc acctccacta 300
ccaacgccga agcctccacg aagacggaaa tccggtcaga aaactgatca aacgactcca 360
tcatcagacg ccgaagcttc atccgatcct gcaccgcctg ttcctgctgc tccagtggct 420

```
cctgttgttc caattgttcc aattgttcca gttcatcctg tacctttgcc aaacggaagt 480
gtaaatactc caaaagtgaa gacgattgca aagacaacag cacgagtact gtattccatt 540
aaacctcaaa taccaccaat tgcgaacaaa actgtgtatc ctgtcaagaa gttgacacct 600
tctgtagttc cgtctccaat gattttaaat ggaaatactg caactgcaag tccatcgaaa 660
aatgcagcat ctgtagttgt cagaaatgca tatacatttt cattacagca aaaagctcca 720
tattatccag ctggaatgcg tccgaaacca actcaaaatg gaattgaaac acctccgaca 780
ggagcacaat ctttaatgcg tgcagccttc tatagtgaga gtcatcctac acgaagtccc 840
ctcgtcccgt atggttttgt tcctccagtt gccacatcat ctacttttgt acctgctgca 900
acaataccat cacctgctag ccgagcaata gctcatcaga agcaaatgct cttaaatacg 960
gaaacgtgtc gacgcgttat gccgtttaat attcaaatgg cattcaaacc acgtcgttgg 1020
gatcctcttc cgaaatcttc tggtgttttg gctcattcaa attcaacaat accatatgtt 1080
cagcgagtcc ccaacaactc aacacaatca gactttcggc cgagaagctt cagtcaaaat 1140
agtgttgctt ccccggctcc tgctcctgtt ccaaatgcaa taaagagaag agaagtaggg 1200
aatttaaaat ctcggcagta tgtgccatgg atcgcaaata gcagggcgtt ggtggcagcg 1260
gcgatggcaa cgatggagga gacggctgaa aagatgtctt cgagcccgtt actatcgtca 1320
caggctccga tgacaacttt gatgccaact cctcctcctc cggctccggc tcctgctcaa 1380
gcatctgcac aatcaacatc agcaactcca gcattggtgg atacaatatc ggctggttca 1440
acgactacag aaacaactac aggcgattca aatcagtcaa atcctccttt acggacgtac 1500
acatcacaca ttcgaaaaac tcctggaaca acactgacac ctgaagaaat cggcgacgcg 1560
attcgaactg aaagtcaaag atttcaagag gatggtgatg aagggccaac agtgaaaagc 1620
ttcctaatga acatctacaa atga                                       1644
```

<210> 48
<211> 547
<212> PRT
<213> Caenorhabditis elegans

<400> 48

```
Leu Pro Glu Ile Glu Gln Glu Leu Lys Ser Glu Ser Val Pro Ala Pro
1               5                   10                  15

Pro Pro Thr Lys Lys Met Ser Ile Ile Ala Asp Ser Trp Asp Asp His
            20                  25                  30

Val Gly Asn Ser Met Glu Glu Glu His Val Asp Glu Lys Asp Ser Glu
        35                  40                  45

Lys Met Glu Asp Ser Glu Gly Arg Gln Asn Val Trp Val Pro Gln Asp
    50                  55                  60

Arg Gly Lys Glu Tyr Ala Pro Glu Gln Tyr Ala Arg Asp Ile Ile Glu
65                  70                  75                  80

His Tyr Ile Pro Ala Ala Arg Asp His Pro Pro Gln Pro Gln Gln Pro
                85                  90                  95

Pro Pro Pro Leu Pro Thr Pro Lys Pro Pro Arg Arg Arg Lys Ser Gly
            100                 105                 110

Gln Lys Thr Asp Gln Thr Thr Pro Ser Ser Asp Ala Glu Ala Ser Ser
        115                 120                 125

Asp Pro Ala Pro Pro Val Pro Ala Ala Pro Val Ala Pro Val Val Pro
    130                 135                 140

Ile Val Pro Ile Val Pro Val His Pro Val Pro Leu Pro Asn Gly Ser
145                 150                 155                 160
```

```
Val Asn Thr Pro Lys Val Lys Thr Ile Ala Lys Thr Thr Ala Arg Val
                165                 170                 175

Leu Tyr Ser Ile Lys Pro Gln Ile Pro Pro Ile Ala Asn Lys Thr Val
            180                 185                 190

Tyr Pro Val Lys Lys Leu Thr Pro Ser Val Val Pro Ser Pro Met Ile
        195                 200                 205

Leu Asn Gly Asn Thr Ala Thr Ala Ser Pro Ser Lys Asn Ala Ala Ser
    210                 215                 220

Val Val Val Arg Asn Ala Tyr Thr Phe Ser Leu Gln Gln Lys Ala Pro
225                 230                 235                 240

Tyr Tyr Pro Ala Gly Met Arg Pro Lys Pro Thr Gln Asn Gly Ile Glu
                245                 250                 255

Thr Pro Pro Thr Gly Ala Gln Ser Leu Met Arg Ala Ala Phe Tyr Ser
            260                 265                 270

Glu Ser His Pro Thr Arg Ser Pro Leu Val Pro Tyr Gly Phe Val Pro
        275                 280                 285

Pro Val Ala Thr Ser Ser Thr Phe Val Pro Ala Ala Thr Ile Pro Ser
    290                 295                 300

Pro Ala Ser Arg Ala Ile Ala His Gln Lys Gln Met Leu Leu Asn Thr
305                 310                 315                 320

Glu Thr Cys Arg Arg Val Met Pro Phe Asn Ile Gln Met Ala Phe Lys
                325                 330                 335

Pro Arg Arg Trp Asp Pro Leu Pro Lys Ser Ser Gly Val Leu Ala His
            340                 345                 350

Ser Asn Ser Thr Ile Pro Tyr Val Gln Arg Val Pro Asn Asn Ser Thr
        355                 360                 365

Gln Ser Asp Phe Arg Pro Arg Ser Phe Ser Gln Asn Ser Val Ala Ser
    370                 375                 380

Pro Ala Pro Ala Pro Val Pro Asn Ala Ile Lys Arg Arg Glu Val Gly
385                 390                 395                 400

Asn Leu Lys Ser Arg Gln Tyr Val Pro Trp Ile Ala Asn Ser Arg Ala
                405                 410                 415

Leu Val Ala Ala Ala Met Ala Thr Met Glu Glu Thr Ala Glu Lys Met
            420                 425                 430

Ser Ser Ser Pro Leu Leu Ser Ser Gln Ala Pro Met Thr Thr Leu Met
        435                 440                 445

Pro Thr Pro Pro Pro Pro Ala Pro Ala Pro Ala Gln Ala Ser Ala Gln
    450                 455                 460

Ser Thr Ser Ala Thr Pro Ala Leu Val Asp Thr Ile Ser Ala Gly Ser
```

465 470 475 480

Thr Thr Thr Glu Thr Thr Thr Gly Asp Ser Asn Gln Ser Asn Pro Pro
485 490 495

Leu Arg Thr Tyr Thr Ser His Ile Arg Lys Thr Pro Gly Thr Thr Leu
500 505 510

Thr Pro Glu Glu Ile Gly Asp Ala Ile Arg Thr Glu Ser Gln Arg Phe
515 520 525

Gln Glu Asp Gly Asp Glu Gly Pro Thr Val Lys Ser Phe Leu Met Asn
530 535 540

Ile Tyr Lys
545

<210> 49
<211> 1248
<212> DNA
<213> Homo sapiens

<400> 49
atgcagatgt cctacgccat ccggtgcgcc ttctaccagc tgctgctggc cgcgctcatg 60
ctggtggcga tgctgcagct gctctacctg tcgctgctgt ccggactgca cgggcaggag 120
gagcaagacc aatattttga gttctttccc ccgtccccac ggtccgtgga ccaggtcaag 180
gcgcagctcc gcaccgcgct ggcctctgga ggcgtcctgg acgctagcgg cgattaccgc 240
gtctacaggg gcctgctgaa gaccaccatg gaccccaacg atgtgatcct ggccacgcac 300
gccagcgtgg acaacctgct gcacctgtcg ggtctgctgg agcgctggga gggcccgctg 360
tccgtgtcgg tgttcgcggc caccaaggag gaggcgcagc tggccacggt gctggcctac 420
gcgctgagca gccactgccc cgacatgcgc gccagggtcg ccatgcacct cgtgtgcccc 480
tcgcgttacg aggcagccgt gcccgacccc cgggagccgg gggagtttgc cctgctgcgg 540
tcctgccagg aggtctttga caagctagcc agggtggccc agcccgggat taattatgcg 600
ctgggcacca atgtctccta ccccaataac ctgctgagga atctggctcg tgaggggggcc 660
aactatgccc tggtgatcga tgtggacatg gtgcccagcg aggggctgtg gagaggcctg 720
cgggaaatgc tggatcagag caaccagtgg ggaggcaccg cgctggtggt gcctgccttc 780
gaaatccgaa gagcccgccg catgcccatg aacaaaaacg agctggtgca gctctaccag 840
gttggcgagg tgcggccctt ctattatggg ttgtgcaccc cctgccaggc acccaccaac 900
tattcccgct gggtcaacct gccggaagag agcttgctgc ggcccgccta cgtggtacct 960
tggcaggacc cctgggagcc attctacgtg gcaggaggca aggtgcccac cttcgacgag 1020
cgctttcggc agtacggctt caaccgaatc agccaggcct gcgagctgca tgtggcgggg 1080
tttgattttg aggtcctgaa cgaaggtttc ttggttcata agggcttcaa agaagcgttg 1140
aagttccatc cccaaaagga ggctgaaaat cagcacaata agatcctata tcgccagttc 1200
aaacaggagt tgaaggccaa gtaccccaac tctccccgac gctgctga 1248

<210> 50
<211> 415
<212> PRT
<213> Homo sapiens

<400> 50
Met Gln Met Ser Tyr Ala Ile Arg Cys Ala Phe Tyr Gln Leu Leu Leu
1 5 10 15

Ala Ala Leu Met Leu Val Ala Met Leu Gln Leu Leu Tyr Leu Ser Leu
20 25 30

Leu Ser Gly Leu His Gly Gln Glu Glu Gln Asp Gln Tyr Phe Glu Phe
35 40 45

Phe Pro Pro Ser Pro Arg Ser Val Asp Gln Val Lys Ala Gln Leu Arg
50 55 60

Thr Ala Leu Ala Ser Gly Gly Val Leu Asp Ala Ser Gly Asp Tyr Arg
65 70 75 80

Val Tyr Arg Gly Leu Leu Lys Thr Thr Met Asp Pro Asn Asp Val Ile
85 90 95

Leu Ala Thr His Ala Ser Val Asp Asn Leu Leu His Leu Ser Gly Leu
100 105 110

Leu Glu Arg Trp Glu Gly Pro Leu Ser Val Ser Val Phe Ala Ala Thr
115 120 125

Lys Glu Glu Ala Gln Leu Ala Thr Val Leu Ala Tyr Ala Leu Ser Ser
130 135 140

His Cys Pro Asp Met Arg Ala Arg Val Ala Met His Leu Val Cys Pro
145 150 155 160

Ser Arg Tyr Glu Ala Ala Val Pro Asp Pro Arg Glu Pro Gly Glu Phe
165 170 175

Ala Leu Leu Arg Ser Cys Gln Glu Val Phe Asp Lys Leu Ala Arg Val
180 185 190

Ala Gln Pro Gly Ile Asn Tyr Ala Leu Gly Thr Asn Val Ser Tyr Pro
195 200 205

Asn Asn Leu Leu Arg Asn Leu Ala Arg Glu Gly Ala Asn Tyr Ala Leu
210 215 220

Val Ile Asp Val Asp Met Val Pro Ser Glu Gly Leu Trp Arg Gly Leu
225 230 235 240

Arg Glu Met Leu Asp Gln Ser Asn Gln Trp Gly Gly Thr Ala Leu Val
245 250 255

Val Pro Ala Phe Glu Ile Arg Arg Ala Arg Arg Met Pro Met Asn Lys
260 265 270

Asn Glu Leu Val Gln Leu Tyr Gln Val Gly Glu Val Arg Pro Phe Tyr
275 280 285

Tyr Gly Leu Cys Thr Pro Cys Gln Ala Pro Thr Asn Tyr Ser Arg Trp
290 295 300

Val Asn Leu Pro Glu Glu Ser Leu Leu Arg Pro Ala Tyr Val Val Pro
305 310 315 320

Trp Gln Asp Pro Trp Glu Pro Phe Tyr Val Ala Gly Gly Lys Val Pro
325 330 335

Thr Phe Asp Glu Arg Phe Arg Gln Tyr Gly Phe Asn Arg Ile Ser Gln

340 345 350

Ala Cys Glu Leu His Val Ala Gly Phe Asp Phe Glu Val Leu Asn Glu
355 360 365

Gly Phe Leu Val His Lys Gly Phe Lys Glu Ala Leu Lys Phe His Pro
370 375 380

Gln Lys Glu Ala Glu Asn Gln His Asn Lys Ile Leu Tyr Arg Gln Phe
385 390 395 400

Lys Gln Glu Leu Lys Ala Lys Tyr Pro Asn Ser Pro Arg Arg Cys
405 410 415

<210> 51
<211> 557
<212> DNA
<213> Homo sapiens

<400> 51
caggcaccca ccaactattc ccgctgggtc aacctgccgg aagagagctt gctgcggccc 60
gcctacgtgg taccttggca ggacccctgg gagccattct acgtggcagg aggcaaggtg 120
cccaccttcg acgagcgctt tcggcagtac ggcttcaacc gaatcaagcc aggcctgcga 180
gctgcatgtg gcggggtttg attttgaggt cctgaacgaa ggtttcttgg ttcataaggg 240
cttcaaagaa gcgttgaagt tccatcccca aaaggaggct gaaaatcagc acaataagat 300
cctatatcgc cagttcaaac aggagttgaa ggccaagtac cccaactctc cccgacgctg 360
ctgagccctt ccctccccta atctgagaag tcagcctctt ggctcctcag gccaccattt 420
aaggcctgac tggggtaaga aatgtcgctn cactttacag angtagcttg tggtgttgaa 480
acactggact tggatatggg gtgcttggga atcgattcct aactttacca ctactaactt 540
gngtggnctt gagtaaa 557

<210> 52
<211> 646
<212> DNA
<213> Homo sapiens

<400> 52
gggtgtgaac acgcatttat ttacacattg tcatcggtag gcacataccc agcccagccc 60
gttgaaggga cagatctcag gctgtcaccc atggnaacct gcagaccctt ccctcctcca 120
gaacccttga aggcgaagcc ctgagagaac caggaatttt aggcttctgt tcaagagcta 180
agaactaaat tttatgcctt catctgattt ctttccaaaa agtccatttc attaagtatt 240
cagacttctt agctccatcc cattcatact ttttgctctc ctactaccca cccaagattg 300
ttaataataa caataataat aacaacaata atactgcgat aatattaata cttcacattt 360
gtacgaagct tacagaatgt tttcacatat agcatctcat ctgagcctcc caacagttcc 420
gtgaggtagg tattctcacc tccctttta cagacagggt aaccgaggct nagagaggta 480
cgggatttac tcaaggccac acagctagtt agtgggtaaa gctaggaatc gatcccagca 540
cccccatatt caaggtccag ggtttaaac accacagctt ccttntggtn aagtgggagc 600
gacatttntt accccagtcc ggcctaaaat gggggcctga nggagc 646

<210> 53
<211> 121
<212> PRT
<213> Homo sapiens

<400> 53
Cys Gln Ala Pro Thr Asn Tyr Ser Arg Trp Val Asn Leu Pro Glu Glu
1 5 10 15

Ser Leu Leu Arg Pro Ala Tyr Val Val Pro Trp Gln Asp Pro Trp Glu
20 25 30

Pro Phe Tyr Val Ala Gly Gly Lys Val Pro Thr Phe Asp Glu Arg Phe
35 40 45

Arg Gln Tyr Gly Phe Asn Arg Ile Ser Gln Ala Cys Glu Leu His Val
50 55 60

Ala Gly Phe Asp Phe Glu Val Leu Asn Glu Gly Phe Leu Val His Lys
65 70 75 80

Gly Phe Lys Glu Ala Leu Lys Phe His Pro Gln Lys Glu Ala Glu Asn
85 90 95

Gln His Asn Lys Ile Leu Tyr Arg Gln Phe Lys Gln Glu Leu Lys Ala
100 105 110

Lys Tyr Pro Asn Ser Pro Arg Arg Cys
115 120

<210> 54
<211> 552
<212> DNA
<213> Homo sapiens

<400> 54
cctccttcct tgcccctccc tacctccctg cctctacatg gcttctctcc acccctccca 60
gagtcctacg ggacaggacc ctgctccagt ggtatccaac tcctccctgc ccactcttcc 120
tcatgggcca cctcacctcc cactttcgat gtctcgcctc ccgtggccac cctgcaatta 180
gctttccaag ccccctcccg tggccgtccc ctcccaagac ctctcaccca tgtagcaatc 240
cctacatggc tgcctgtcat gtccctactc tctaagccct cctgcccact gttcctccct 300
ccccgacatg ctgacaccaa gtggtggaaa ccacccctca gccccaagcc tgccctgtgc 360
agagttcagc tttgtgttga atgagggggg agagggacaa gtgagggcgg agagagaagt 420
tcaggaggag gcagggatgc gcanggagca ganagtgaag gaaggaagat ccgaacagat 480
tgacngaaaa cgttggaccg naaaagttgg tttntcctaa tttttttccg ggagaacccg 540
ttacacagtt nt 552

<210> 55
<211> 754
<212> DNA
<213> Homo sapiens

<400> 55
ggaaagtttt tgtttttcttt ttcccacaca tttccggggt tggggtgttt tccaagactc 60
agacacattt gttaacaaag agagaaaaaa aactggggga gcagggagcc cgtgggcaaa 120
gaagtgaccc cagcagtctg tggacaatgc cttgctccct cttccctgct gaccgcgccc 180
agcgggtgcc acaggctgct gctcgtggaa tctagagtat ttgtctgtaa tatatatctg 240
catttgcctt tcccctccct cccaccccnn acccctgctc ctccagcagc ttccccatca 300
atgcacgtcg nccggcggna cacacagaac aggccttccg tcaggcctga gccccttccc 360
tgggcggcac caaagcaggt gccnntnctg gtgaggggag ttggggcact tgccccagcc 420
nancanactn acacctgggc cantncggna nnccctntnn cnttcnntcn aaccnattct 480
ggaancccnn ngggaannaa nnggnancnt annnccnnna tcncnaannn aanaatnanc 540
nannnncnng nnncnnncnn ncannnccan ncaccccang nnacgnnana tcgnannccc 600
ttgnnctcaa ancgaancgc cnccacnncc tacagganca nanncnnaac tcagngaaan 660
tcnaccntac tnccanncan cncttcccaa ccnnntcatc cttcnttncn ncnatccnnc 720

aatataannn cncttataca naactccacn ngnt 754

<210> 56
<211> 555
<212> DNA
<213> Homo sapiens

<400> 56
gggcagccgg gcgagtccca gggcaggggt cctccgccgc cttgcacctg ccctgctggg 60
cggcaccggg tcagtgccct gccccctcct gcgggtccca actctctctt tcccatcgtg 120
cgtcctctgg agaagtgcgc gcgtgagctg acatggaccc aaatcctcgg gccgccctgg 180
agcgccagca gctccgcctt cgggagcggc aaaaattctt cgaggacatt ttacaagcca 240
gagacagagt ttgtctttcc tctgtcccat ctgcatctcg agtcgcagag accccccata 300
ggtaagtatc tcatccatgg aagtgaatgt ggacacactg gagcaagtag aacttattga 360
ccttggggac ccggatgcag cagatgtgtt cttgccttgc gaagatcctt caccaacccc 420
caagtccgtc tgggatggac aaccatttgg aaggaagctg agcctgcccg tgcctacatc 480
agacaaggac cacattanga acttcttctt cctcctcctn cgactcttca ncaanctgga 540
taagccaaat caagt 555

<210> 57
<211> 611
<212> DNA
<213> Homo sapiens

<400> 57
nggcntnttn ttnttagttn nttnttnanc ccggnctttn nnccaaaagg gnnntanggg 60
ggcncnnngg ntngngnccc ncnttacttg annngtnang ggntnnncat ccntttaaaa 120
ccncttgnnt gaaaaactgt tttaaaaaac tncggnangg ttnaggggng ggaanagnnc 180
taaaaaaagc nggggntttt ngnccaaccn aantttntnt tncctaattn gcaaatcntn 240
tntcaggggt aanccaaaaa ctggnggnag gnttncnccn ggaaaaantt accnttaaan 300
cagganaggg ttaaattntn aaaaagggcc ccaattcccc ccatcnttcc caccttnggg 360
ggccnnctgc nagtaaanag nctggtcttt tccccaanag ggnttttggc tggcccngng 420
gcccnnattn gggnnnaatn cccccnccgn gggcacaann nttncaagcc agggcccccc 480
nttggttaaa tttnaagggn nccnagggtt tttgncccct ttnanaaccc ccctttnccc 540
cnccttttaa aanccnncct ttccccccaa nggnnccttt ttntgncccc aanannnacc 600
tgggatttgg g 611

<210> 58
<211> 4425
<212> DNA
<213> Homo sapiens

<400> 58
atggggaaga acaaactcct tcatccaagt ctggttcttc tcctcttggt cctcctgccc 60
acagacgcct cagtctctgg aaaaccgcag tatatggttc tggtcccctc cctgctccac 120
actgagacca ctgagaaggg ctgtgtcctt ctgagctacc tgaatgagac agtgactgta 180
agtgcttcct tggagtctgt caggggaaac aggagcctct tcactgacct ggaggcggag 240
aatgacgtac tccactgtgt cgccttcgct gtcccaaagt cttcatccaa tgaggaggta 300
atgttcctca ctgtccaagt gaaaggacca acccaagaat ttaagaagcg gaccacagtg 360
atggttaaga acgaggacag tctggtcttt gtccagacag acaaatcaat ctacaaacca 420
gggcagacag tgaaatttcg tgttgtctcc atggatgaaa actttcaccc cctgaatgag 480
ttgattccac tagtatacat tcaggatccc aaaggaaatc gcatcgcaca atggcagagt 540
ttccagttag agggtggcct caagcaattt tcttttcccc tctcatcaga gcccttccag 600
ggctcctaca aggtggtggt acagaagaaa tcaggtggaa ggacagagca ccctttcacc 660
gtggaggaat ttgttcttcc caagtttgaa gtacaagtaa cagtgccaaa gataatcacc 720
atcttggaag aagagatgaa tgtatcagtg tgtggcctat acacatatgg gaagcctgtc 780
cctggacatg tgactgtgag catttgcaga aagtatagtg acgcttccga ctgccacggt 840
gaagattcac aggctttctg tgagaaattc agtggacagc taaacagcca tggctgcttc 900

tatcagcaag taaaaaccaa ggtcttccag ctgaagagga aggagtatga aatgaaactt 960
cacactgagg cccagatcca agaagaagga acagtggtgg aattgactgg aaggcagtcc 1020
agtgaaatca caagaaccat aaccaaactc tcatttgtga aagtggactc acactttcga 1080
cagggaattc ccttctttgg gcaggtgcgc ctagtagatg ggaaaggcgt ccctatacca 1140
aataaagtca tattcatcag aggaaatgaa gcaaactatt actccaatgc taccacggat 1200
gagcatggcc ttgtacagtt ctctatcaac accaccaacg ttatgggtac ctctcttact 1260
gttagggtca attacaagga tcgtagtccc tgttacggct accagtgggt gtcagaagaa 1320
cacgaagagg cacatcacac tgcttatctt gtgttctccc caagcaagag ctttgtccac 1380
cttgagccca tgtctcatga actaccctgt ggccatactc agacagtcca ggcacattat 1440
attctgaatg gaggcaccct gctggggctg aagaagctct cctttttatta tctgataatg 1500
gcaaagggag gcattgtccg aactgggact catggactgc ttgtgaagca ggaagacatg 1560
aagggccatt tttccatctc aatccctgtg aagtcagaca ttgctcctgt cgctcggttg 1620
ctcatctatg ctgttttacc taccggggac gtgattgggg attctgcaaa atatgatgtt 1680
gaaaattgtc tggccaacaa ggtggatttg agcttcagcc catcacaaag tctcccagcc 1740
tcacacgccc acctgcgagt cacagcggct cctcagtccg tctgcgccct ccgtgctgtg 1800
gaccaaagcg tgctgctcat gaagcctgat gctgagctct cggcgtcctc ggtttacaac 1860
ctgctaccag aaaaggacct cactggcttc cctgggcctt tgaatgacca ggacgatgaa 1920
gactgcatca atcgtcataa tgtctatatt aatggaatca catatactcc agtatcaagt 1980
acaaatgaaa aggatatgta cagcttccta gaggacatgg gcttaaaggc attcaccaac 2040
tcaaagattc gtaaacccaa aatgtgtcca cagcttcaac agtatgaaat gcatggacct 2100
gaaggtctac gtgtaggttt ttatgagtca gatgtaatgg gaagaggcca tgcacgcctg 2160
gtgcatgttg aagagcctca cacggagacc gtacgaaagt acttccctga gacatggatc 2220
tgggatttgg tggtggtaaa ctcagcaggg gtggctgagg taggagtaac agtccctgac 2280
accatcaccg agtggaaggc aggggccttc tgcctgtctg aagatgctgg acttggtatc 2340
tcttccactg cctctctccg agccttccag cccttctttg tggagcttac aatgccttac 2400
tctgtgattc gtggagaggc cttcacactc aaggccacgg tcctaaacta ccttcccaaa 2460
tgcatccggg tcagtgtgca gctggaagcc tctcccgcct tccttgctgt cccagtggag 2520
aaggaacaag cgcctcactg catctgtgca aacgggcggc aaactgtgtc ctgggcagta 2580
accccaaagt cattaggaaa tgtgaatttc actgtgagcg cagaggcact agagtctcaa 2640
gagctgtgtg ggactgaggt gccttcagtt cctgaacacg gaaggaaaga cacagtcatc 2700
aagcctctgt tggttgaacc tgaaggacta gagaaggaaa caacattcaa ctccctactt 2760
tgtccatcag gtggtgaggt ttctgaagaa ttatccctga aactgccacc aaatgtggta 2820
gaagaatctg cccgagcttc tgtctcagtt ttgggagaca tattaggctc tgccatgcaa 2880
aacacacaaa atcttctcca gatgccctat ggctgtggag agcagaatat ggtcctcttt 2940
gctcctaaca tctatgtact ggattatcta aatgaaacac agcagcttac tccagaggtc 3000
aagtccaagg ccattggcta tctcaacact ggttaccaga gacagttgaa ctacaaacac 3060
tatgatggct cctacagcac ctttggggag cgatatggca ggaaccaggg caacacctgg 3120
ctcacagcct ttgttctgaa gacttttgcc caagctcgag cctacatctt catcgatgaa 3180
gcacacatta cccaagccct catatggctc tcccagaggc agaaggacaa tggctgtttc 3240
aggagctctg ggtcactgct caacaatgcc ataaagggag gagtagaaga tgaagtgacc 3300
ctctccgcct atatcaccat cgcccttctg gagattcctc tcacagtcac tcaccctgtt 3360
gtccgcaatg ccctgttttg cctggagtca gcctggaaga cagcacaaga aggggaccat 3420
ggcagccatg tatataccaa agcactgctg gcctatgctt ttgccctggc aggtaaccag 3480
gacaagagga aggaagtact caagtcactt aatgaggaag ctgtgaagaa agacaactct 3540
gtccattggg agcgccctca gaaacccaag gcaccagtgg ggcattttta cgaaccccag 3600
gctccctctg ctgaggtgga gatgacatcc tatgtgctcc tcgcttatct cacggcccag 3660
ccagccccaa cctcggagga cctgacctct gcaaccaaca tcgtgaagtg gatcacgaag 3720
cagcagaatg cccagggcgg tttctcctcc acccaggaca cagtggtggc tctccatgct 3780
ctgtccaaat atggagccgc cacatttacc aggactggga aggctgcaca ggtgactatc 3840
cagtcttcag ggacattttc cagcaaattc caagtggaca acaacaatcg cctgttactg 3900
cagcaggtct cattgccaga gctgcctggg gaatacagca tgaaagtgac aggagaagga 3960
tgtgtctacc tccagacctc cttgaaatac aatattctcc cagaaaagga agagttcccc 4020
tttgctttag gagtgcagac tctgcctcaa acttgtgatg aacccaaagc ccacaccagc 4080
ttccaaatct ccctaagtgt cagttacaca gggagccgct ctgcctccaa catggcgatc 4140
gttgatgtga agatggtctc tggcttcatt cccctgaagc caacagtgaa aatgcttgaa 4200
agatctaacc atgtgagccg gacagaagtc agcagcaacc atgtcttgat ttaccttgat 4260
aaggtgtcaa atcagacact gagcttgttc ttcacggttc tgcaagatgt cccagtaaga 4320
gatctcaaac cagccatagt gaaagtctat gattactacg agacggatga gtttgcaatc 4380

gctgagtaca atgctccttg cagcaaagat cttggaaatg cttga 4425

<210> 59
<211> 1474
<212> PRT
<213> Homo sapiens

<400> 59
Met Gly Lys Asn Lys Leu Leu His Pro Ser Leu Val Leu Leu Leu Leu
1 5 10 15

Val Leu Leu Pro Thr Asp Ala Ser Val Ser Gly Lys Pro Gln Tyr Met
20 25 30

Val Leu Val Pro Ser Leu Leu His Thr Glu Thr Thr Glu Lys Gly Cys
35 40 45

Val Leu Leu Ser Tyr Leu Asn Glu Thr Val Thr Val Ser Ala Ser Leu
50 55 60

Glu Ser Val Arg Gly Asn Arg Ser Leu Phe Thr Asp Leu Glu Ala Glu
65 70 75 80

Asn Asp Val Leu His Cys Val Ala Phe Ala Val Pro Lys Ser Ser Ser
85 90 95

Asn Glu Glu Val Met Phe Leu Thr Val Gln Val Lys Gly Pro Thr Gln
100 105 110

Glu Phe Lys Lys Arg Thr Thr Val Met Val Lys Asn Glu Asp Ser Leu
115 120 125

Val Phe Val Gln Thr Asp Lys Ser Ile Tyr Lys Pro Gly Gln Thr Val
130 135 140

Lys Phe Arg Val Val Ser Met Asp Glu Asn Phe His Pro Leu Asn Glu
145 150 155 160

Leu Ile Pro Leu Val Tyr Ile Gln Asp Pro Lys Gly Asn Arg Ile Ala
165 170 175

Gln Trp Gln Ser Phe Gln Leu Glu Gly Gly Leu Lys Gln Phe Ser Phe
180 185 190

Pro Leu Ser Ser Glu Pro Phe Gln Gly Ser Tyr Lys Val Val Val Gln
195 200 205

Lys Lys Ser Gly Gly Arg Thr Glu His Pro Phe Thr Val Glu Glu Phe
210 215 220

Val Leu Pro Lys Phe Glu Val Gln Val Thr Val Pro Lys Ile Ile Thr
225 230 235 240

Ile Leu Glu Glu Glu Met Asn Val Ser Val Cys Gly Leu Tyr Thr Tyr
245 250 255

Gly Lys Pro Val Pro Gly His Val Thr Val Ser Ile Cys Arg Lys Tyr
260 265 270

```
Ser Asp Ala Ser Asp Cys His Gly Glu Asp Ser Gln Ala Phe Cys Glu
        275                 280                 285

Lys Phe Ser Gly Gln Leu Asn Ser His Gly Cys Phe Tyr Gln Gln Val
    290                 295                 300

Lys Thr Lys Val Phe Gln Leu Lys Arg Lys Glu Tyr Glu Met Lys Leu
305                 310                 315                 320

His Thr Glu Ala Gln Ile Gln Glu Glu Gly Thr Val Val Glu Leu Thr
                325                 330                 335

Gly Arg Gln Ser Ser Glu Ile Thr Arg Thr Ile Thr Lys Leu Ser Phe
            340                 345                 350

Val Lys Val Asp Ser His Phe Arg Gln Gly Ile Pro Phe Phe Gly Gln
        355                 360                 365

Val Arg Leu Val Asp Gly Lys Gly Val Pro Ile Pro Asn Lys Val Ile
    370                 375                 380

Phe Ile Arg Gly Asn Glu Ala Asn Tyr Tyr Ser Asn Ala Thr Thr Asp
385                 390                 395                 400

Glu His Gly Leu Val Gln Phe Ser Ile Asn Thr Thr Asn Val Met Gly
                405                 410                 415

Thr Ser Leu Thr Val Arg Val Asn Tyr Lys Asp Arg Ser Pro Cys Tyr
            420                 425                 430

Gly Tyr Gln Trp Val Ser Glu Glu His Glu Glu Ala His His Thr Ala
        435                 440                 445

Tyr Leu Val Phe Ser Pro Ser Lys Ser Phe Val His Leu Glu Pro Met
    450                 455                 460

Ser His Glu Leu Pro Cys Gly His Thr Gln Thr Val Gln Ala His Tyr
465                 470                 475                 480

Ile Leu Asn Gly Gly Thr Leu Leu Gly Leu Lys Lys Leu Ser Phe Tyr
                485                 490                 495

Tyr Leu Ile Met Ala Lys Gly Gly Ile Val Arg Thr Gly Thr His Gly
            500                 505                 510

Leu Leu Val Lys Gln Glu Asp Met Lys Gly His Phe Ser Ile Ser Ile
        515                 520                 525

Pro Val Lys Ser Asp Ile Ala Pro Val Ala Arg Leu Leu Ile Tyr Ala
    530                 535                 540

Val Leu Pro Thr Gly Asp Val Ile Gly Asp Ser Ala Lys Tyr Asp Val
545                 550                 555                 560

Glu Asn Cys Leu Ala Asn Lys Val Asp Leu Ser Phe Ser Pro Ser Gln
                565                 570                 575
```

Ser Leu Pro Ala Ser His Ala His Leu Arg Val Thr Ala Ala Pro Gln
580 585 590

Ser Val Cys Ala Leu Arg Ala Val Asp Gln Ser Val Leu Leu Met Lys
595 600 605

Pro Asp Ala Glu Leu Ser Ala Ser Ser Val Tyr Asn Leu Leu Pro Glu
610 615 620

Lys Asp Leu Thr Gly Phe Pro Gly Pro Leu Asn Asp Gln Asp Asp Glu
625 630 635 640

Asp Cys Ile Asn Arg His Asn Val Tyr Ile Asn Gly Ile Thr Tyr Thr
645 650 655

Pro Val Ser Ser Thr Asn Glu Lys Asp Met Tyr Ser Phe Leu Glu Asp
660 665 670

Met Gly Leu Lys Ala Phe Thr Asn Ser Lys Ile Arg Lys Pro Lys Met
675 680 685

Cys Pro Gln Leu Gln Gln Tyr Glu Met His Gly Pro Glu Gly Leu Arg
690 695 700

Val Gly Phe Tyr Glu Ser Asp Val Met Gly Arg Gly His Ala Arg Leu
705 710 715 720

Val His Val Glu Glu Pro His Thr Glu Thr Val Arg Lys Tyr Phe Pro
725 730 735

Glu Thr Trp Ile Trp Asp Leu Val Val Val Asn Ser Ala Gly Val Ala
740 745 750

Glu Val Gly Val Thr Val Pro Asp Thr Ile Thr Glu Trp Lys Ala Gly
755 760 765

Ala Phe Cys Leu Ser Glu Asp Ala Gly Leu Gly Ile Ser Ser Thr Ala
770 775 780

Ser Leu Arg Ala Phe Gln Pro Phe Phe Val Glu Leu Thr Met Pro Tyr
785 790 795 800

Ser Val Ile Arg Gly Glu Ala Phe Thr Leu Lys Ala Thr Val Leu Asn
805 810 815

Tyr Leu Pro Lys Cys Ile Arg Val Ser Val Gln Leu Glu Ala Ser Pro
820 825 830

Ala Phe Leu Ala Val Pro Val Glu Lys Glu Gln Ala Pro His Cys Ile
835 840 845

Cys Ala Asn Gly Arg Gln Thr Val Ser Trp Ala Val Thr Pro Lys Ser
850 855 860

Leu Gly Asn Val Asn Phe Thr Val Ser Ala Glu Ala Leu Glu Ser Gln
865 870 875 880

Glu Leu Cys Gly Thr Glu Val Pro Ser Val Pro Glu His Gly Arg Lys

885 890 895

Asp Thr Val Ile Lys Pro Leu Leu Val Glu Pro Glu Gly Leu Glu Lys
900 905 910

Glu Thr Thr Phe Asn Ser Leu Leu Cys Pro Ser Gly Gly Glu Val Ser
915 920 925

Glu Glu Leu Ser Leu Lys Leu Pro Pro Asn Val Val Glu Glu Ser Ala
930 935 940

Arg Ala Ser Val Ser Val Leu Gly Asp Ile Leu Gly Ser Ala Met Gln
945 950 955 960

Asn Thr Gln Asn Leu Leu Gln Met Pro Tyr Gly Cys Gly Glu Gln Asn
965 970 975

Met Val Leu Phe Ala Pro Asn Ile Tyr Val Leu Asp Tyr Leu Asn Glu
980 985 990

Thr Gln Gln Leu Thr Pro Glu Val Lys Ser Lys Ala Ile Gly Tyr Leu
995 1000 1005

Asn Thr Gly Tyr Gln Arg Gln Leu Asn Tyr Lys His Tyr Asp Gly Ser
1010 1015 1020

Tyr Ser Thr Phe Gly Glu Arg Tyr Gly Arg Asn Gln Gly Asn Thr Trp
1025 1030 1035 1040

Leu Thr Ala Phe Val Leu Lys Thr Phe Ala Gln Ala Arg Ala Tyr Ile
1045 1050 1055

Phe Ile Asp Glu Ala His Ile Thr Gln Ala Leu Ile Trp Leu Ser Gln
1060 1065 1070

Arg Gln Lys Asp Asn Gly Cys Phe Arg Ser Ser Gly Ser Leu Leu Asn
1075 1080 1085

Asn Ala Ile Lys Gly Gly Val Glu Asp Glu Val Thr Leu Ser Ala Tyr
1090 1095 1100

Ile Thr Ile Ala Leu Leu Glu Ile Pro Leu Thr Val Thr His Pro Val
1105 1110 1115 1120

Val Arg Asn Ala Leu Phe Cys Leu Glu Ser Ala Trp Lys Thr Ala Gln
1125 1130 1135

Glu Gly Asp His Gly Ser His Val Tyr Thr Lys Ala Leu Leu Ala Tyr
1140 1145 1150

Ala Phe Ala Leu Ala Gly Asn Gln Asp Lys Arg Lys Glu Val Leu Lys
1155 1160 1165

Ser Leu Asn Glu Glu Ala Val Lys Lys Asp Asn Ser Val His Trp Glu
1170 1175 1180

Arg Pro Gln Lys Pro Lys Ala Pro Val Gly His Phe Tyr Glu Pro Gln
1185 1190 1195 1200

```
Ala Pro Ser Ala Glu Val Glu Met Thr Ser Tyr Val Leu Leu Ala Tyr
                1205                1210                1215

Leu Thr Ala Gln Pro Ala Pro Thr Ser Glu Asp Leu Thr Ser Ala Thr
            1220                1225                1230

Asn Ile Val Lys Trp Ile Thr Lys Gln Gln Asn Ala Gln Gly Gly Phe
        1235                1240                1245

Ser Ser Thr Gln Asp Thr Val Val Ala Leu His Ala Leu Ser Lys Tyr
    1250                1255                1260

Gly Ala Ala Thr Phe Thr Arg Thr Gly Lys Ala Ala Gln Val Thr Ile
1265                1270                1275                1280

Gln Ser Ser Gly Thr Phe Ser Ser Lys Phe Gln Val Asp Asn Asn Asn
                1285                1290                1295

Arg Leu Leu Leu Gln Gln Val Ser Leu Pro Glu Leu Pro Gly Glu Tyr
            1300                1305                1310

Ser Met Lys Val Thr Gly Glu Gly Cys Val Tyr Leu Gln Thr Ser Leu
        1315                1320                1325

Lys Tyr Asn Ile Leu Pro Glu Lys Glu Glu Phe Pro Phe Ala Leu Gly
    1330                1335                1340

Val Gln Thr Leu Pro Gln Thr Cys Asp Glu Pro Lys Ala His Thr Ser
1345                1350                1355                1360

Phe Gln Ile Ser Leu Ser Val Ser Tyr Thr Gly Ser Arg Ser Ala Ser
                1365                1370                1375

Asn Met Ala Ile Val Asp Val Lys Met Val Ser Gly Phe Ile Pro Leu
            1380                1385                1390

Lys Pro Thr Val Lys Met Leu Glu Arg Ser Asn His Val Ser Arg Thr
        1395                1400                1405

Glu Val Ser Ser Asn His Val Leu Ile Tyr Leu Asp Lys Val Ser Asn
    1410                1415                1420

Gln Thr Leu Ser Leu Phe Phe Thr Val Leu Gln Asp Val Pro Val Arg
1425                1430                1435                1440

Asp Leu Lys Pro Ala Ile Val Lys Val Tyr Asp Tyr Tyr Glu Thr Asp
                1445                1450                1455

Glu Phe Ala Ile Ala Glu Tyr Asn Ala Pro Cys Ser Lys Asp Leu Gly
            1460                1465                1470

Asn Ala


<210> 60
<211> 722
```

<212> DNA
<213> Homo sapiens

<400> 60
ngaccnnaaa aagngtttat tcatcaagtn tttaaagatn caaaaacacg tgtnttntgn 60
ggagctntga naacaggact ccagcaaagc acttttcagc cttggggnct tcaagcattt 120
ccaagatctt tgctgcaagg agcattgnnc tcagcaattg caaactcatc cttntcgtag 180
naancataga ctttcactat ggctggtttc ananctntta ctgggacatt ttgcanaacc 240
gngaanaaca agctcagggn ctgatttgac accttntcaa ggtaaatcaa gacatggttg 300
ctgntgactt ctgncccggn tcacatgggt tagatctttc aagcnttttt nactgnnngg 360
cttcagggga atgaaacccc gagaccntnt tnncaatnaa cgacncccnt nttgggaggc 420
aaaccggntc cctgngtaac ctnnccctta ggggantatt ggaanctng gtgtgggncn 480
tttgggttca tnnnnaaggt ttngaggcna agnnctgnct tcnnaaagca aaggggnacc 540
tnttccnttt ttntggnnaa antttgnttt ttcaaggnat tnnngaagnt annnncaacc 600
ttctcccggn nntttcaang cnggntttcc cagggnagtt ttggnatagn nccnnttnna 660
aaanncgggg ggttttttac ncccttggaa ttntnnggga aaaanncctn aaannnggga 720
ac 722

<210> 61
<211> 557
<212> DNA
<213> Homo sapiens

<400> 61
gtgtgcagtg ccccttgcc tcttcaaacc tacagcttct ctttgccatt tgtggatttc 60
acatcactcc acacagaaac attacagcct ggcatcccca gtctttgcct tcttccagct 120
gcctcgacac agcactgtgg cctgtcccta ttgcccaggc acgccatttc caagggcagg 180
aaggggcagt gtcctgaagc ccatcttttc tgtgactgtc ttaggtgatg tgtagccccc 240
tccacctttc cactcaacaa cctcccaccc ctgtcctgct gcatggtccg gagtctggga 300
cctactttgt tttttgttat ttatgacctt gtttaaagaa aataaatatc tcccaacctt 360
taaaaaaaaa aaaaaaaaaa aactcgagag atctatgaat cgaagatact gaaaaacccc 420
gcangttcac ttcaactgtg catcgtgcan catctcaatt ctttcatttn atacatccnt 480
tttgcctttc tttatgtaac tatactcctc taaagtttca atcttgggca ttnaaccttt 540
gatctataaa attttta 557

<210> 62
<211> 640
<212> DNA
<213> Homo sapiens

<400> 62
ttaaaggttg ggagatattt attttcttta aacaaggtca taaataacaa aaaacaaagt 60
aggtcccaga ctccggacca tgcagcagga caggggtggg aggttgttga gtggaaaggt 120
ggaggggct acacatcacc taagacagtc acagaaaaga tgggcttcag gacactgccc 180
cttcctgccc ttggaaatgg cgtgcctggg caatagggac aggccacagt gctgtgtcga 240
ggcagctgga agaaggcaaa gactggggat gccaggctgt aatgtttctg tgtggagtga 300
tgtgaaatcc acaaatggca aagagaagct gtaggtttga agaggcaagg gggcactgca 360
cacgtcgacg cggccgcgaa ttcggatccc cggggcctcc atggccatat gaccacccaa 420
gctagcgtaa tctggaacat cgtatgggta aagccataga gatctctttt tttgggtttg 480
gtggggtatc ttcatcatcg aatagatagt tatatacata tccattgtag tgggattaaa 540
catccctgta gtgattccaa acgcgttata cgcagtttgg tccgtccaac caggtgacag 600
nggtttgaat tattaccatc tcaattttac tagccgggat 640

<210> 63
<211> 566
<212> DNA
<213> Homo sapiens

<400> 63
ctcccccatc cgcatcctcc tcctcctgtc ccctcttcct ctctgtccct caccaagcca 60
tcccctcttt gcctctcccc atccctcgct ggcctgctcc tgctcccctt ctcccatcct 120
cccctccccc gtctctgccc agccagcccc ctcccgactc cccagtttca tcggactccc 180
tggccccatc ccgtccccgc cctggcccct ttgtgcccct tctcatcgtt ttctccctcc 240
ttcccgggct tggcgtccct tctccccccct aactccttcc ctcggcctcc ctgccccttc 300
acggcccgcc tgcctccctt gcccaagtcc tgagccacca tgctgacccc gatggtggcc 360
cggnggggtg gtgtccccgg actcttctct ntncaagaac acgcttcagc cggctgcccc 420
aagctacgct gggaggaggc cgacgcagcn ttgcctnagc caggcctggt ggtcctttgn 480
ccagncatca tggccttcan tggcggtnca tnngnttcac ctnctggnag cacntattga 540
taccaanatg ggtggccnta tnncta 566

<210> 64
<211> 648
<212> DNA
<213> Homo sapiens

<400> 64
tggcgttcac gcagatcgtc ttttattagc ggtctgtaaa gcacctccca gggtcccccg 60
accccagatt ggaggaagcc ttgagaggtc agtagcacca gggacatggc agggcccgag 120
ggcgcgatgt gcagccgatg gtgagggact gggcgccctc gcctgccccc ggggttgtca 180
gcactgggaa ggcttggggg tagcagccac ctccttcccc caacccaaca gactagttca 240
aatttgggta aataaataaa ataaataaga ttcctcaagc tggcctaccc tggagaggag 300
ccgtggttgc agccggccac tcgggaggcc cgagggccag cgggggttag ttgggcgtc 360
ctctcctctc gggtgatggg gagccctggg ggatggcagc atagggggctg ggatggcctt 420
ggcagaggcc gtctncccac attctgactc cttggtcccc cttgaaaccc tgttggtgtc 480
ccttcccaca aagcccttct tgccctcagt gggtggggaa ggccggtgcc cccttccctt 540
cttcancgca aagggtntgc aggaaagggg caaaattagg gggnaaaaag gtccctttt 600
tcancccccnt tngtccccna aaagatgggg ccttttccnt ttngnggt 648

<210> 65
<211> 2274
<212> PRT
<213> Mus sp.

<400> 65
Met Thr Ser Ser Met Ala Ser Tyr Glu Gln Leu Val Arg Gln Val Glu
1 5 10 15

Ala Leu Lys Ala Glu Asn Thr His Leu Arg Gln Glu Leu Arg Asp Asn
20 25 30

Ser Ser His Leu Ser Lys Leu Glu Thr Glu Thr Ser Gly Met Lys Glu
35 40 45

Val Leu Lys His Leu Gln Gly Lys Leu Glu Gln Glu Ala Arg Val Leu
50 55 60

Val Ser Ser Gly Gln Thr Glu Val Leu Glu Gln Leu Lys Ala Leu Gln
65 70 75 80

Thr Asp Ile Ser Ser Leu Tyr Asn Leu Lys Phe His Ala Pro Ala Leu
85 90 95

Gly Pro Glu Pro Ala Ala Arg Thr Pro Glu Gly Ser Pro Val His Gly
100 105 110

Ser Gly Pro Ser Lys Asp Ser Phe Gly Glu Leu Ser Arg Ala Thr Ile

115 120 125

Arg Leu Leu Glu Glu Leu Asp Gln Glu Arg Cys Phe Leu Leu Ser Glu
130 135 140

Ile Glu Lys Glu Glu Lys Glu Lys Leu Trp Tyr Tyr Ser Gln Leu Gln
145 150 155 160

Gly Leu Ser Lys Arg Leu Asp Glu Leu Pro His Val Asp Thr Phe Ser
165 170 175

Met Gln Met Asp Leu Ile Arg Gln Gln Leu Glu Phe Glu Ala Gln His
180 185 190

Ile Arg Ser Leu Met Glu Glu Arg Phe Gly Thr Ser Asp Glu Met Val
195 200 205

Gln Arg Ala Gln Ile Arg Ala Ser Arg Leu Glu Gln Ile Asp Lys Glu
210 215 220

Leu Leu Glu Ala Gln Asp Arg Val Gln Gln Thr Glu Pro Gln Ala Leu
225 230 235 240

Leu Ala Val Lys Pro Val Ala Val Glu Glu Glu Gln Glu Ala Glu Val
245 250 255

Pro Thr His Pro Glu Asp Gly Thr Pro Gln Pro Gly Asn Ser Lys Val
260 265 270

Glu Val Val Phe Trp Leu Leu Ser Met Leu Ala Thr Arg Asp Gln Glu
275 280 285

Asp Thr Ala Arg Thr Leu Leu Ala Met Ser Ser Ser Pro Glu Ser Cys
290 295 300

Val Ala Met Arg Arg Ser Gly Cys Leu Pro Leu Leu Leu Gln Ile Leu
305 310 315 320

His Gly Thr Glu Ala Gly Ser Val Gly Arg Ala Gly Ile Pro Gly Ala
325 330 335

Pro Gly Ala Lys Asp Ala Arg Met Arg Ala Asn Ala Ala Leu His Asn
340 345 350

Ile Val Phe Ser Gln Pro Asp Gln Gly Leu Ala Arg Lys Glu Met Arg
355 360 365

Val Leu His Val Leu Glu Gln Ile Arg Ala Tyr Cys Glu Thr Cys Trp
370 375 380

Asp Trp Leu Gln Ala Arg Asp Ser Gly Thr Glu Thr Pro Val Pro Ile
385 390 395 400

Glu Pro Gln Ile Cys Gln Ala Thr Cys Ala Val Met Lys Leu Ser Phe
405 410 415

Asp Glu Glu Tyr Arg Arg Ala Met Asn Glu Leu Gly Gly Leu Gln Ala
420 425 430

Val Ala Glu Leu Leu Gln Val Asp Tyr Glu Met His Lys Met Thr Arg
435 440 445

Asp Pro Leu Asn Leu Ala Leu Arg Arg Tyr Ala Gly Met Thr Leu Thr
450 455 460

Asn Leu Thr Phe Gly Asp Val Ala Asn Lys Ala Thr Leu Cys Ala Arg
465 470 475 480

Arg Gly Cys Met Glu Ala Ile Val Ala Gln Leu Gly Ser Glu Ser Glu
485 490 495

Glu Leu His Gln Val Val Ser Ser Ile Leu Arg Asn Leu Ser Trp Arg
500 505 510

Ala Asp Ile Asn Ser Lys Lys Val Leu Arg Glu Val Gly Ser Met Thr
515 520 525

Ala Leu Met Glu Cys Val Leu Arg Ala Ser Lys Glu Ser Thr Leu Lys
530 535 540

Ser Val Leu Ser Ala Leu Trp Asn Leu Ser Ala His Ser Thr Glu Asn
545 550 555 560

Lys Ala Ala Ile Cys Gln Val Asp Gly Ala Leu Gly Phe Leu Val Ser
565 570 575

Thr Leu Thr Tyr Arg Cys Gln Gly Asn Ser Leu Ala Val Ile Glu Ser
580 585 590

Gly Gly Gly Ile Leu Arg Asn Val Ser Ser Leu Ile Ala Thr Arg Glu
595 600 605

Asp Tyr Arg Gln Val Leu Arg Asp His Asn Cys Leu Gln Thr Leu Leu
610 615 620

Gln His Leu Thr Ser His Ser Leu Thr Ile Val Ser Asn Ala Cys Gly
625 630 635 640

Thr Leu Trp Asn Leu Ser Ala Arg Ser Pro Arg Asp Gln Glu Leu Leu
645 650 655

Trp Asp Leu Gly Ala Val Gly Met Leu Arg Asn Leu Val His Ser Lys
660 665 670

His Lys Met Ile Ala Met Gly Ser Ala Ala Ala Leu Arg Asn Leu Leu
675 680 685

Ala His Arg Pro Ala Lys Tyr Gln Ala Ala Ala Met Ala Val Ser Pro
690 695 700

Gly Thr Cys Val Pro Ser Leu Tyr Val Arg Lys Gln Arg Ala Leu Glu
705 710 715 720

Ala Glu Leu Asp Thr Arg His Leu Val His Ala Leu Gly His Leu Glu
725 730 735

Lys Gln Ser Leu Pro Glu Ala Glu Thr Thr Ser Lys Lys Pro Leu Pro
740 745 750

Pro Leu Arg His Leu Asp Gly Leu Val Gln Asp Tyr Ala Ser Asp Ser
755 760 765

Gly Cys Phe Asp Asp Asp Asp Ala Pro Ser Leu Ala Ala Ala Ala Thr
770 775 780

Thr Ala Glu Pro Ala Ser Pro Ala Val Met Ser Met Phe Leu Gly Gly
785 790 795 800

Pro Phe Leu Gln Gly Gln Ala Leu Ala Arg Thr Pro Pro Ala Arg Gln
805 810 815

Gly Gly Leu Glu Ala Glu Lys Glu Ala Gly Gly Glu Ala Ala Val Ala
820 825 830

Ala Lys Ala Lys Ala Lys Leu Ala Leu Ala Val Ala Arg Ile Asp Arg
835 840 845

Leu Val Glu Asp Ile Ser Ala Leu His Thr Ser Ser Asp Asp Ser Phe
850 855 860

Ser Leu Ser Ser Gly Asp Pro Gly Gln Glu Ala Pro Arg Glu Gly Arg
865 870 875 880

Ala Gln Ser Cys Ser Pro Cys Arg Gly Thr Glu Gly Gly Arg Arg Glu
885 890 895

Ala Gly Ser Arg Ala His Pro Leu Leu Arg Leu Lys Ala Ala His Thr
900 905 910

Ser Leu Ser Asn Asp Ser Leu Asn Ser Gly Ser Thr Ser Asp Gly Tyr
915 920 925

Cys Thr Arg Glu His Met Thr Pro Cys Pro Leu Ala Ala Leu Ala Glu
930 935 940

His Arg Asp Asp Pro Val Arg Gly Gln Thr Arg Pro Arg Arg Leu Asp
945 950 955 960

Leu Asp Leu Pro Ser Arg Ala Glu Leu Pro Ala Arg Asp Thr Ala Ala
965 970 975

Thr Asp Ala Arg Val Arg Thr Ile Lys Leu Ser Pro Thr Tyr Gln His
980 985 990

Val Pro Leu Leu Asp Gly Ala Ala Gly Ala Gly Val Arg Pro Leu Val
995 1000 1005

Gly Pro Gly Thr Ser Pro Gly Ala Arg Lys Gln Ala Trp Ile Pro Ala
1010 1015 1020

Asp Ser Leu Ser Lys Val Pro Glu Lys Leu Val Ala Ser Pro Leu Pro
1025 1030 1035 1040

Ile Ala Ser Lys Val Leu Gln Lys Leu Val Ala Gln Asp Gly Pro Met

1045 1050 1055

Ser Leu Ser Arg Cys Ser Ser Leu Ser Ser Leu Ser Ser Thr Gly His
1060 1065 1070

Ala Val Pro Ser Gln Ala Glu Asn Leu Asp Ser Asp Ser Ser Leu Glu
1075 1080 1085

Gly Leu Glu Glu Ala Gly Pro Gly Glu Ala Glu Leu Gly Arg Ala Trp
1090 1095 1100

Arg Ala Ser Gly Ser Thr Ser Leu Pro Val Ser Ile Pro Ala Pro Gln
1105 1110 1115 1120

Arg Gly Arg Ser Arg Gly Leu Gly Val Glu Asp Ala Thr Pro Ser Ser
1125 1130 1135

Ser Ser Glu Asn Cys Val Gln Glu Thr Pro Leu Val Leu Ser Arg Cys
1140 1145 1150

Ser Ser Val Ser Ser Leu Gly Ser Phe Glu Ser Arg Ser Ile Ala Ser
1155 1160 1165

Ser Ile Pro Ser Asp Pro Cys Ser Gly Leu Gly Ser Gly Thr Val Ser
1170 1175 1180

Pro Ser Glu Leu Pro Asp Ser Pro Gly Gln Thr Met Pro Pro Ser Arg
1185 1190 1195 1200

Ser Lys Thr Pro Pro Ala Pro Pro Gly Gln Pro Glu Thr Ser Gln Phe
1205 1210 1215

Ser Leu Gln Trp Glu Ser Tyr Val Lys Arg Phe Leu Asp Ile Ala Asp
1220 1225 1230

Cys Arg Glu Arg Cys Gln Pro Pro Ser Glu Leu Asp Ala Gly Ser Val
1235 1240 1245

Arg Phe Thr Val Glu Lys Pro Asp Glu Asn Phe Ser Cys Ala Ser Ser
1250 1255 1260

Leu Ser Ala Leu Ala Leu His Glu Leu Tyr Val Gln Gln Asp Val Glu
1265 1270 1275 1280

Leu Arg Leu Arg Pro Pro Ala Cys Pro Glu Arg Ala Val Gly Gly Gly
1285 1290 1295

Gly His Arg Arg Arg Asp Glu Ala Ala Ser Arg Leu Asp Gly Pro Ala
1300 1305 1310

Pro Ala Gly Ser Arg Ala Arg Ser Ala Thr Asp Lys Glu Leu Glu Ala
1315 1320 1325

Leu Arg Glu Cys Leu Gly Ala Ala Met Pro Ala Arg Leu Arg Lys Val
1330 1335 1340

Ala Ser Ala Leu Val Pro Gly Arg Arg Ser Leu Pro Val Pro Val Tyr
1345 1350 1355 1360

Met Leu Val Pro Ala Pro Ala Arg Gly Asp Asp Ser Gly Thr Asp Ser
1365 1370 1375

Ala Glu Gly Thr Pro Val Asn Phe Ser Ser Ala Ala Ser Leu Ser Asp
1380 1385 1390

Glu Thr Leu Gln Gly Pro Ser Arg Asp Lys Pro Ala Gly Pro Gly Asp
1395 1400 1405

Arg Gln Lys Pro Thr Gly Arg Ala Ala Pro Ala Arg Gln Thr Arg Ser
1410 1415 1420

His Arg Pro Lys Ala Ala Gly Ala Gly Lys Ser Thr Glu His Thr Arg
1425 1430 1435 1440

Gly Pro Cys Arg Asn Arg Ala Gly Leu Glu Leu Pro Leu Ser Arg Pro
1445 1450 1455

Gln Ser Ala Arg Ser Asn Arg Asp Ser Ser Cys Gln Thr Arg Thr Arg
1460 1465 1470

Gly Asp Gly Ala Leu Gln Ser Leu Cys Leu Thr Thr Pro Thr Glu Glu
1475 1480 1485

Ala Val Tyr Cys Phe Tyr Asp Ser Asp Glu Glu Pro Pro Ala Thr Ala
1490 1495 1500

Pro Pro Pro Arg Arg Ala Ser Ala Ile Pro Arg Ala Leu Lys Arg Glu
1505 1510 1515 1520

Lys Pro Ala Gly Arg Lys Glu Thr Pro Ser Arg Ala Ala Gln Pro Ala
1525 1530 1535

Thr Leu Pro Val Arg Ala Gln Pro Arg Leu Ile Val Asp Glu Thr Pro
1540 1545 1550

Pro Cys Tyr Ser Leu Thr Ser Ser Ala Ser Ser Leu Ser Glu Pro Glu
1555 1560 1565

Ala Pro Glu Gln Pro Ala Asn His Ala Arg Gly Pro Glu Gln Gly Ser
1570 1575 1580

Lys Gln Asp Ser Ser Pro Ser Pro Arg Ala Glu Glu Glu Leu Leu Gln
1585 1590 1595 1600

Arg Cys Ile Ser Leu Ala Met Pro Arg Arg Arg Thr Gln Val Pro Gly
1605 1610 1615

Ser Arg Arg Arg Lys Pro Arg Ala Leu Arg Ser Asp Ile Arg Pro Thr
1620 1625 1630

Glu Ile Thr Gln Lys Cys Gln Glu Glu Val Ala Gly Ser Asp Pro Ala
1635 1640 1645

Ser Asp Leu Asp Ser Val Glu Trp Gln Ala Ile Gln Glu Gly Ala Asn
1650 1655 1660

Ser Ile Val Thr Trp Leu His Gln Ala Ala Ala Lys Ala Ser Leu Glu
1665 1670 1675 1680

Ala Ser Ser Glu Ser Asp Ser Leu Leu Ser Leu Val Ser Gly Val Ser
1685 1690 1695

Ala Gly Ser Thr Leu Gln Pro Ser Lys Leu Arg Lys Gly Arg Lys Pro
1700 1705 1710

Ala Ala Glu Ala Gly Gly Ala Trp Arg Pro Glu Lys Arg Gly Thr Thr
1715 1720 1725

Ser Thr Lys Ile Asn Gly Ser Pro Arg Leu Pro Asn Gly Pro Glu Lys
1730 1735 1740

Ala Lys Gly Thr Gln Lys Met Met Ala Gly Glu Ser Thr Met Leu Arg
1745 1750 1755 1760

Gly Arg Thr Val Ile Tyr Ser Ala Gly Pro Ala Ser Arg Thr Gln Ser
1765 1770 1775

Lys Gly Ile Ser Gly Pro Cys Thr Thr Pro Lys Lys Thr Gly Thr Ser
1780 1785 1790

Gly Thr Thr Gln Pro Glu Thr Val Thr Lys Ala Pro Ser Pro Glu Gln
1795 1800 1805

Gln Arg Ser Arg Ser Leu His Arg Pro Gly Lys Ile Ser Glu Leu Ala
1810 1815 1820

Ala Leu Arg His Pro Pro Arg Ser Ala Thr Pro Pro Ala Arg Leu Ala
1825 1830 1835 1840

Lys Thr Pro Ser Ser Ser Ser Ser Gln Thr Ser Pro Ala Ser Gln Pro
1845 1850 1855

Leu Pro Arg Arg Ser Pro Leu Ala Thr Pro Thr Gly Gly Pro Leu Pro
1860 1865 1870

Gly Pro Gly Gly Ser Leu Val Pro Lys Ser Pro Ala Arg Ala Leu Leu
1875 1880 1885

Ala Lys Gln His Lys Thr Gln Lys Ser Pro Val Arg Ile Pro Phe Met
1890 1895 1900

Gln Arg Pro Ala Arg Arg Val Pro Pro Pro Leu Ala Arg Pro Ser Pro
1905 1910 1915 1920

Glu Pro Gly Ser Arg Gly Arg Ala Gly Ala Glu Gly Thr Pro Gly Ala
1925 1930 1935

Arg Gly Ser Arg Leu Gly Leu Val Arg Met Ala Ser Ala Arg Ser Ser
1940 1945 1950

Gly Ser Glu Ser Ser Asp Arg Ser Gly Phe Arg Arg Gln Leu Thr Phe
1955 1960 1965

Ile Lys Glu Ser Pro Gly Leu Leu Arg Arg Arg Arg Ser Glu Leu Ser

1970 1975 1980

Ser Ala Asp Ser Thr Ala Ser Thr Ser Gln Ala Ala Ser Pro Arg Arg
1985 1990 1995 2000

Gly Arg Pro Ala Leu Pro Ala Val Phe Leu Cys Ser Ser Arg Cys Asp
2005 2010 2015

Glu Leu Arg Val Ser Pro Arg Gln Pro Leu Ala Ala Gln Arg Ser Pro
2020 2025 2030

Gln Ala Lys Pro Gly Leu Ala Pro Leu Ala Pro Arg Arg Thr Ser Ser
2035 2040 2045

Glu Ser Pro Ser Arg Leu Pro Val Arg Ala Ser Pro Gly Arg Pro Glu
2050 2055 2060

Thr Val Lys Arg Tyr Ala Ser Leu Pro His Ile Ser Val Ser Arg Arg
2065 2070 2075 2080

Ser Asp Ser Ala Val Ser Val Pro Thr Thr Gln Ala Asn Ala Thr Arg
2085 2090 2095

Arg Gly Ser Asp Gly Glu Ala Arg Pro Leu Pro Arg Val Ala Pro Pro
2100 2105 2110

Gly Thr Thr Trp Arg Arg Ile Lys Asp Glu Asp Val Pro His Ile Leu
2115 2120 2125

Arg Ser Thr Leu Pro Ala Thr Ala Leu Pro Leu Arg Val Ser Ser Pro
2130 2135 2140

Glu Asp Ser Pro Ala Gly Thr Pro Gln Arg Lys Thr Ser Asp Ala Val
2145 2150 2155 2160

Val Gln Thr Glu Asp Val Ala Thr Ser Lys Thr Asn Ser Ser Thr Ser
2165 2170 2175

Pro Ser Leu Glu Ser Arg Asp Pro Pro Gln Ala Pro Ala Ser Gly Pro
2180 2185 2190

Val Ala Pro Gln Gly Ser Asp Val Asp Gly Pro Val Leu Thr Lys Pro
2195 2200 2205

Pro Ala Ser Ala Pro Phe Pro His Glu Gly Leu Ser Ala Val Ile Ala
2210 2215 2220

Gly Phe Pro Thr Ser Arg His Gly Ser Pro Ser Arg Ala Ala Arg Val
2225 2230 2235 2240

Pro Pro Phe Asn Tyr Val Pro Ser Pro Met Ala Ala Ala Thr Met Ala
2245 2250 2255

Ser Asp Ser Ala Val Glu Lys Ala Pro Val Ser Ser Pro Ala Ser Leu
2260 2265 2270

Leu Glu

<210> 66
<211> 942
<212> DNA
<213> Caenorhabditis elegans

<400> 66
atgtacgaag atcagtactg cgtggggtat aatttctgg aaggaacagg agactttcga 60
gaagacggtc tagaacctat caccctggca tcccatgcca cttcagatat gatgttaact 120
ctggaaaaaa tgacctcaat gtgggatgga cctatatcag ttgggatatt tattgatttt 180
cactctagtc aagctctgga gtatctcgca gaagtgcaca gatgtgatga ggagttcagg 240
aagaagatga caattcactt tgcaatccgt cagtcagcat tccaacaaac ttgcccaaaa 300
attcaaattc cagcttcaga cagaacttgc tggaagttca gagcggatca atcctacctc 360
cgaagccatc tgtcaggacc cttccaacta tatccgagca accttatgag aaatttggct 420
cgccagggag ccaagtcgga tattcatttt attatggatg cagatatgat tgttagtgag 480
ggattcgcgc ggaaactcaa aaaagtggca aatgagatga tcgacggaaa aagtaaaaaa 540
gtattggcaa ttcgaagatt cgaatcggtt aatggaactt atttgcctag aactcacttt 600
gagttgaagc aatctatggc ttattccaac ggatatgaat gggaagttca agtaattctt 660
caccgtaacg acccctataa cgcggcctat ttcccgtcca gaatcaaagt tatgcactct 720
ctgatttacg ccctttgtcg tgctggttac acgtttcatg ttccgtcaca cgttttcgat 780
gttcatgaag ggattaagca tactaataca atttattcga aagccacaat tgctcatcag 840
gaagcttatg cgatggacat agccggagcc agatatgtca gagaaatgga cgaaaagtac 900
ccggacactt tggacaagtg tggaagattc aagatgtatt ag 942

<210> 67
<211> 313
<212> PRT
<213> Caenorhabditis elegans

<400> 67
Met Tyr Glu Asp Gln Tyr Cys Val Gly Tyr Asn Phe Leu Glu Gly Thr
1 5 10 15

Gly Asp Phe Arg Glu Asp Gly Leu Glu Pro Ile Thr Leu Ala Ser His
20 25 30

Ala Thr Ser Asp Met Met Leu Thr Leu Glu Lys Met Thr Ser Met Trp
35 40 45

Asp Gly Pro Ile Ser Val Gly Ile Phe Ile Asp Phe His Ser Ser Gln
50 55 60

Ala Leu Glu Tyr Leu Ala Glu Val His Arg Cys Asp Glu Glu Phe Arg
65 70 75 80

Lys Lys Met Thr Ile His Phe Ala Ile Arg Gln Ser Ala Phe Gln Gln
85 90 95

Thr Cys Pro Lys Ile Gln Ile Pro Ala Ser Asp Arg Thr Cys Trp Lys
100 105 110

Phe Arg Ala Asp Gln Ser Tyr Leu Arg Ser His Leu Ser Gly Pro Phe
115 120 125

Gln Leu Tyr Pro Ser Asn Leu Met Arg Asn Leu Ala Arg Gln Gly Ala
130 135 140

```
Lys Ser Asp Ile His Phe Ile Met Asp Ala Asp Met Ile Val Ser Glu
145                 150                 155                 160

Gly Phe Ala Arg Lys Leu Lys Lys Val Ala Asn Glu Met Ile Asp Gly
                165                 170                 175

Lys Ser Lys Lys Val Leu Ala Ile Arg Arg Phe Glu Ser Val Asn Gly
            180                 185                 190

Thr Tyr Leu Pro Arg Thr His Phe Glu Leu Lys Gln Ser Met Ala Tyr
        195                 200                 205

Ser Asn Gly Tyr Glu Trp Glu Val Gln Val Ile Leu His Arg Asn Asp
    210                 215                 220

Pro Tyr Asn Ala Ala Tyr Phe Pro Ser Arg Ile Lys Val Met His Ser
225                 230                 235                 240

Leu Ile Tyr Ala Leu Cys Arg Ala Gly Tyr Thr Phe His Val Pro Ser
                245                 250                 255

His Val Phe Asp Val His Glu Gly Ile Lys His Thr Asn Thr Ile Tyr
            260                 265                 270

Ser Lys Ala Thr Ile Ala His Gln Glu Ala Tyr Ala Met Asp Ile Ala
        275                 280                 285

Gly Ala Arg Tyr Val Arg Glu Met Asp Glu Lys Tyr Pro Asp Thr Leu
    290                 295                 300

Asp Lys Cys Gly Arg Phe Lys Met Tyr
305                 310
```

<210> 68
<211> 4527
<212> DNA
<213> Caenorhabditis elegans

<400> 68

```
atgcggctgc taattctcaa tatattattt gtggtttggc aaatacatgg cgtcatcgga 60
cagagtacaa atgcagctgt ggtgtcaaca accgcggcgc cagttaaacc agccacttat 120
atgcttgtcg ccccagcagt agtgcgaccc gatcaaccgt tctccgtttg catgaacctt 180
ctgaaacaag ctacggatga agacatgatt gttcggattg aagtgagaac tgaacggaac 240
gagacgattg ctgctcgagt tataagcaat ttgaagccgg gaatcgcaca gactgtcagc 300
ttgagtgaga tgcctgctca atcactaact cctaggcaaa gctataagct ctacatccgc 360
ggggagactc ttaacgcaga actcatattt gaaaacgaaa atgagttaaa gtacgatcaa 420
aaagctttat cagttttcat ccaaacagat cgagcaatct atcggcctgc ttctcttgtt 480
cgataccggg caattgtagt gaaatcggac ttgaaacctt acgtcggaaa cgctactatt 540
aaaatttcg atccaagcag aaatctgatt tcacagacga ttggtgtaac tttggataga 600
ggagtataca gcggcgagct ccaacttgcc gaagagactc tccttggaga ttggtttatc 660
gaggtggaaa cctcgaatgg agtccaggac aagtcgtcat ttacggtaga tacatacgtg 720
cttcccaaat ttgaagtgaa tatcaagacg tcttcgttta taactattaa cgatgatttg 780
tcagtgttcg tcgatgcaaa gtacacatat ggaaaaggcg ttgctggaaa ggccaaagta 840
tctctagagc ttccatggca cagatggcat gcgatggttc ccactatcat tgacgaaaat 900
ggagtcaaga aggaagagga gctgatggtc gagcgtactg taaagctgaa cagacaagga 960
gaagcggcag ttgttttctc gaatgatgag ctcaagaggc acaagttgct acatgaatgg 1020
ggaggaggtt caattcgaat tgttgcctca gtcactgaag acatcactga aatcgaaaga 1080
```

aacgctactc atcaaatttc aacgttccgt gaggaggtca aacttgatgt tgaaaagcag 1140
ggagacacct tcaagccagg actcacttat aatgttgtag ttgcactgaa acaaatggat 1200
gacactcccg tgaaggcaac tttgccaaag agagttcagg tatcaacatt ctacaactat 1260
ccatacaatc atgacacttc atctctgcaa gaggaaaaag aaacaaaaat tgtggaagtt 1320
gacgctcacg gcacctcggt tctcaccctc caaccaccaa tcaactgcac ttctgctcgc 1380
attgaggctc attatgatat tggtggcaag gacaacttca cagccacccc aatatattcc 1440
tcgctatatg tcgaggcagc cgtgagccca acgaagagtt tcctgcagct tttggctgat 1500
aatgagggag ctgttgatgt aggaaagagt cttagcttct cgttgaaggc tacccagccg 1560
ctgtcgacaa ttacctatca ggtcatgtct cgttcgaaca ttgtggtgtc tcaacaaatg 1620
actgttaatt ccgaacatgc gaccattagc ttcccagcaa cggctaacat ggccccaaag 1680
tcacggttga tcgtctatgc cattattgag agtagtcaag aagttcttgt tgatgcactt 1740
gactttaaag tagaaggaat attccaaaac caagtggcac tatcaattga caaacaagca 1800
gtcgagccag gacagaatgt gaagtttaag gtcacatctg acaagaactc ttttgttggg 1860
cttcttgtgg tagatcagag tgttctcttg ttaaaaactg gcaatgatat cactagagaa 1920
aaagttgagc aagacttgga gaattatgac tctaataatg ttggaggagg attcgggggt 1980
ccacggccat gggaggcaat tgatagaaag aagaggtcca tttggagacc ttggtgggga 2040
attggaggga gtgatgctca atcaatcttt tctaatgctg ggcttgtagt tttgaccgac 2100
gccctcttgt accgcgagcc tcaacgtgaa tttatgtcag taatgatgat ggatggtgct 2160
ccaggaatgg ctgaagccgc attcgccgct ccgccaatgg gcgggtcgtc gcctccaccg 2220
ccgacagttc gaaagttctt cccacacacc tggatttggt cggatctcaa ctctactagt 2280
ggagaagttg agatggaaat tgaagctccg gacaccatca cctcgtgggt cgcctccact 2340
tttgcaatca atgaggaaaa tggtctcggt gtggctccaa cgacttcaaa acttcgcgtg 2400
tttcgtccat tttttattca actcaatctt ccatatgctg tccgtcgtgg agaaaagttt 2460
gcacttttgg tgcttgtttt caattatatg gaaaaggagc aagatgtaac agtgactctg 2520
aagtatgata aagactctgg ttacgatctc ctgaagaaag atggaactgt agtaaggcgt 2580
gacgaggttg gccaacaaaa tgttcgtatt gtgtccgttg ctggtggagg aacatcgaaa 2640
gcagtgtact tcccaattgt gccgtcttca attggagaga tcccggtaca tataagcgct 2700
attgcatccc aaggtggtga tgcagtagag atgaacttga gagttgatcc acaaggatat 2760
aaagtagata gaaatattcc atttgtcatt gacctcaaca ataactcttc ggatttctcc 2820
aaaaatcttg agcttatctg gccaaacgac gtggttgatg gatcgcagaa ggcaagatta 2880
gatgtgattg gtgatatgat gggaccagtg ttaaacaatg ctcataagct cgtacaaatg 2940
ccatatggat gtggagaaca gaatatgttg aacttggttc caaatatttt ggtggtcaag 3000
tacttaagag caacaaatag aaatgagagc caattagaga caaaggctat caagtttatt 3060
gagcagggaa ttcagaggga gttgacgtac aaacgtgctg acaactcttt ctcagctttt 3120
ggcgatagtg ataaagctgg ttcaacgtgg ctcaccgcat ttgtcgttcg atcgttccat 3180
cacgccaaac aatacgcatt cgtggatcct aatgttattt cccgagcagt cgcattcttg 3240
aactctcaac aaatggagtc tggtgcattt gccgaacgtg gtgaagttca tcacaaggat 3300
atgcagggag gtgctcagga cgggggtgta gctctaacgg catttgtgct aatctcgatt 3360
ttggagaatg gaatggagaa tggaaaggcg gtgacgtatt tggaaaagca tttggatgaa 3420
gtatctggaa atgcgtatac aatggctgta gtggcatatg cacttcaatt ggcaaaaagt 3480
aagcaagctg gaaaggcatt cgagaatttg aagaaacata agatcgttga gaaaagtgga 3540
gatgtgaagt ttgccagtgc tcagaaaaaa gtggagaaac taaaagaatc gagagcgtat 3600
atgttccagg cccgtccagt agacattgaa actacctctt acgcagttct ttcttacctc 3660
gcccaaaatc aaacctcaga atctctctca attattcgtt ggttagtatc ccaacgcaac 3720
gagctaggag gattcacttc cactcaagac actgtcatgg ctcttcaagc tctctcttcc 3780
tacgcagcag tcacttattc cgacaaacac acaagtcaag taacaattct caacggaaaa 3840
catactcact catttgatat caacataaga aacgcaattg ttctccaatc ctatcaacta 3900
tcttcattga atgatgcagt ttcaattaat gccaatggaa ctggtgtggt cttcgctcag 3960
ttgagttatt cttactacag ggactctttg aatgacgatg ctccgttctt ctgctcccag 4020
gagatcaagg aaataagagc tggaaatcga ctgcaattgg atttgtgttg caactatacc 4080
cgtcctggaa agtctaacat ggctctagcc gaaattgatg ccttgagtgg ttaccgattc 4140
gatgcggaac aggtgcatac tttgacaagt attgaggatt tgcaaagggt ggagatggaa 4200
aaggacgaca ctaagatgaa cgtatatttc aacccgctcg gtggccgtcc ggtatgtctc 4260
tcgttatact ctgatgtcac ttatcaagtt gccgatcaaa aacctgccaa cttccgatta 4320
gtggattact acgatccgga ggagcagctc aaaatgacat acgccgccaa gcaaacacga 4380
tcgctgcaag agaaatgtgg ggaagactgt tggccaccga tttccccatc actgcctcca 4440
tttgacgagt cgacagtaac aggcacgtcc tctggttttg gtgccaagtg gtgtgctctt 4500
atcattgctg tgctccttat tgcatga 4527

<210> 69
<211> 4560
<212> DNA
<213> Caenorhabditis elegans

<400> 69
atgcggctgc taattctcaa tatattattt gtggtttggc aaatacatgg cgtcatcgga 60
cagagtacaa atgcagctgt ggtgtcaaca accgcggcgc cagttaaacc agccacttat 120
atgcttgtcg ccccagcagt agtgcgaccc gatcaaccgt tctccgtttg catgaacctt 180
ctgaaacaag ctacggatga agacatgatt gttcggattg aagtgagaac tgaacggaac 240
gagacgattg ctgctcgagt tataagcaat ttgaagccgg gaatcgcaca gactgtcagc 300
ttgagtgaga tgcctgctca atcactaact cctaggcaaa gctataagct ctacatccgc 360
ggggagactc ttaacgcaga actcatattt gaaaacgaaa atgagttaaa gtacgatcaa 420
aaagctttat cagttttcat ccaaacagat cgagcaatct atcggcctgc ttctcttgtt 480
cgataccggg caattgtagt gaaatcggac ttgaaacctt acgtcggaaa cgctactatt 540
aaaattttcg atccaagcag aaatctgatt tcacagacga ttggtgtaac tttggataga 600
ggagtataca gcggcgagct ccaacttgcc gaagagactc tccttggaga ttggtttatc 660
gaggtggaaa cctcgaatgg agtccaggac aagtcgtcat ttacggtaga tacatacgtg 720
cttcccaaat ttgaagtgaa tatcaagacg tcttcgttta taactattaa cgatgatttg 780
tcagtgttcg tcgatgcaaa gtacacatat ggaaaaggcg ttgctggaaa ggccaaagta 840
tctctagagc ttccatggca cagatggcat gcgatggttc ccactatcat tgacgaaaat 900
ggagtcaaga aggaagagga gctgatggtc gagcgtactg taaagctgaa cagacaagga 960
gaagcggcag ttgttttctc gaatgatgag ctcaagaggc acaagttgct acatgaatgg 1020
ggaggaggtt caattcgaat tgttgcctca gtcactgaag acatcactga aatcgaaaga 1080
aacgctactc atcaaatttc aacgttccgt gaggaggtca aacttgatgt tgaaaagcag 1140
ggagacacct tcaagccagg actcacttat aatgttgtag ttgcactgaa acaaatggat 1200
gacactcccg tgaaggcaac tttgccaaag agagttcagg tatcaacatt ctacaactat 1260
ccatacaatc atgacacttc atctctgcaa gaggaaaaag aaacaaaaat tgtggaagtt 1320
gacgctcacg gcacctcggt tctcaccctc caaccaccaa tcaactgcac ttctgctcgc 1380
attgaggctc attatgatat tggtggcaag gacaacttca cagccacccc aatatattcc 1440
tcgctatatg tcgaggcagc cgtgagccca acgaagagtt tcctgcagct tttggctgat 1500
aatgagggag ctgttgatgt aggaaagagt cttagcttct cgttgaaggc tacccagccg 1560
ctgtcgacaa ttacctatca ggtcatgtct cgttcgaaca ttgtggtgtc tcaacaaatg 1620
actgttaatt ccgaacatgc gaccattagc ttcccagcaa cggctaacat ggccccaaag 1680
tcacggttga tcgtctatgc cattattgag agtagtcaag aagttcttgt tgatgcactt 1740
gactttaaag tagaaggaat attccaaaac caagtggcac tatcaattga caaacaagca 1800
gtcgagccag gacagaatgt gaagtttaag gtcacatctg acaagaactc ttttgttggg 1860
cttcttgtgg tagatcagag tgttctcttg ttaaaaactg gcaatgatat cactagagaa 1920
aaagttgagc aagacttgga gaattatgac tctaataatg ttggaggagg attcgggggt 1980
ccacggccat gggaggcaat tgatagaaag aagaggtcca tttggagacc ttggtgggga 2040
attggaggga gtgatgctca atcaatcttt tctaatgctg ggcttgtagt tttgaccgac 2100
gccctcttgt accgcgagcc tcaacgtgaa tttatgtcag aaagaagact taatacaccc 2160
ggaggtctta cagtaatgat gatggatggt gctccaggaa tggctgaagc cgcattcgcc 2220
gctccgccaa tgggcgggtc gtcgcctcca ccgccgacag ttcgaaagtt cttcccacac 2280
acctggattt ggtcggatct caactctact agtggagaag ttgagatgga aattgaagct 2340
ccggacacca tcacctcgtg ggtcgcctcc acttttgcaa tcaatgagga aaatggtctc 2400
ggtgtggctc caacgacttc aaaacttcgc gtgtttcgtc catttttttat tcaactcaat 2460
cttccatatg ctgtccgtcg tggagaaaag tttgcacttt tggtgcttgt tttcaattat 2520
atggaaaagg agcaagatgt aacagtgact ctgaagtatg ataaagactc tggttacgat 2580
ctcctgaaga aagatggaac tgtagtaagg cgtgacgagg ttggccaaca aaatgttcgt 2640
attgtgtccg ttgctggtgg aggaacatcg aaagcagtgt acttcccaat tgtgccgtct 2700
tcaattggag agatcccggt acatataagc gctattgcat cccaaggtgg tgatgcagta 2760
gagatgaact tgagagttga tccacaagga tataaagtag atagaaatat tccatttgtc 2820
attgacctca acaataactc ttcggatttc tccaaaaatc ttgagcttat ctggccaaac 2880
gacgtggttg atggatcgca gaaggcaaga ttagatgtga ttggtgatat gatgggacca 2940
gtgttaaaca atgctcataa gctcgtacaa atgccatatg gatgtggaga acagaatatg 3000
ttgaacttgg ttccaaatat tttggtggtc aagtacttaa gagcaacaaa tagaaatgag 3060

```
agccaattag agacaaaggc tatcaagttt attgagcagg gaattcagag ggagttgacg 3120
tacaaacgtg ctgacaactc tttctcagct tttggcgata gtgataaagc tggttcaacg 3180
tggctcaccg catttgtcgt tcgatcgttc catcacgcca aacaatacgc attcgtggat 3240
cctaatgtta tttcccgagc agtcgcattc ttgaactctc aacaaatgga gtctggtgca 3300
tttgccgaac gtggtgaagt tcatcacaag gatatgcagg gaggtgctca ggacgggggt 3360
gtagctctaa cggcatttgt gctaatctcg attttggaga atggaatgga gaatggaaag 3420
gcggtgacgt atttggaaaa gcatttggat gaagtatctg gaaatgcgta tacaatggct 3480
gtagtggcat atgcacttca attggcaaaa agtaagcaag ctggaaaggc attcgagaat 3540
ttgaagaaac ataagatcgt tgagaaaagt ggagatgtga agtttgccag tgctcagaaa 3600
aaagtggaga aactaaaaga atcgagagcg tatatgttcc aggcccgtcc agtagacatt 3660
gaaactacct cttacgcagt tctttcttac ctcgcccaaa atcaaacctc agaatctctc 3720
tcaattattc gttggttagt atcccaacgc aacgagctag gaggattcac ttccactcaa 3780
gacactgtca tggctcttca agctctctct tcctacgcag cagtcactta ttccgacaaa 3840
cacacaagtc aagtaacaat tctcaacgga aaacatactc actcatttga tatcaacata 3900
agaaacgcaa ttgttctcca atcctatcaa ctatcttcat tgaatgatgc agtttcaatt 3960
aatgccaatg gaactggtgt ggtcttcgct cagttgagtt attcttacta cagggactct 4020
ttgaatgacg atgctccgtt cttctgctcc caggagatca aggaaataag agctggaaat 4080
cgactgcaat tggatttgtg ttgcaactat acccgtcctg gaaagtctaa catggctcta 4140
gccgaaattg atgccttgag tggttaccga ttcgatgcgg aacaggtgca tactttgaca 4200
agtattgagg atttgcaaag ggtggagatg gaaaaggacg acactaagat gaacgtatat 4260
ttcaacccgc tcggtggccg tccggtatgt ctctcgttat actctgatgt cacttatcaa 4320
gttgccgatc aaaaacctgc caacttccga ttagtggatt actacgatcc ggaggagcag 4380
ctcaaaatga catacgccgc caagcaaaca cgatcgctgc aagagaaatg tggggaagac 4440
tgttggccac cgatttcccc atcactgcct ccatttgacg agtcgacagt aacaggcacg 4500
tcctctggtt ttggtgccaa gtggtgtgct cttatcattg ctgtgctcct tattgcatga 4560
```

<210> 70
<211> 1508
<212> PRT
<213> Caenorhabditis elegans

<400> 70

```
Met Arg Leu Leu Ile Leu Asn Ile Leu Phe Val Val Trp Gln Ile His
1               5                   10                  15

Gly Val Ile Gly Gln Ser Thr Asn Ala Ala Val Val Ser Thr Thr Ala
            20                  25                  30

Ala Pro Val Lys Pro Ala Thr Tyr Met Leu Val Ala Pro Ala Val Val
        35                  40                  45

Arg Pro Asp Gln Pro Phe Ser Val Cys Met Asn Leu Leu Lys Gln Ala
    50                  55                  60

Thr Asp Glu Asp Met Ile Val Arg Ile Glu Val Arg Thr Glu Arg Asn
65                  70                  75                  80

Glu Thr Ile Ala Ala Arg Val Ile Ser Asn Leu Lys Pro Gly Ile Ala
                85                  90                  95

Gln Thr Val Ser Leu Ser Glu Met Pro Ala Gln Ser Leu Thr Pro Arg
            100                 105                 110

Gln Ser Tyr Lys Leu Tyr Ile Arg Gly Glu Thr Leu Asn Ala Glu Leu
        115                 120                 125

Ile Phe Glu Asn Glu Asn Glu Leu Lys Tyr Asp Gln Lys Ala Leu Ser
    130                 135                 140
```

```
Val Phe Ile Gln Thr Asp Arg Ala Ile Tyr Arg Pro Ala Ser Leu Val
145                 150                 155                 160

Arg Tyr Arg Ala Ile Val Val Lys Ser Asp Leu Lys Pro Tyr Val Gly
                165                 170                 175

Asn Ala Thr Ile Lys Ile Phe Asp Pro Ser Arg Asn Leu Ile Ser Gln
            180                 185                 190

Thr Ile Gly Val Thr Leu Asp Arg Gly Val Tyr Ser Gly Glu Leu Gln
        195                 200                 205

Leu Ala Glu Glu Thr Leu Leu Gly Asp Trp Phe Ile Glu Val Glu Thr
    210                 215                 220

Ser Asn Gly Val Gln Asp Lys Ser Ser Phe Thr Val Asp Thr Tyr Val
225                 230                 235                 240

Leu Pro Lys Phe Glu Val Asn Ile Lys Thr Ser Ser Phe Ile Thr Ile
                245                 250                 255

Asn Asp Asp Leu Ser Val Phe Val Asp Ala Lys Tyr Thr Tyr Gly Lys
            260                 265                 270

Gly Val Ala Gly Lys Ala Lys Val Ser Leu Glu Leu Pro Trp His Arg
        275                 280                 285

Trp His Ala Met Val Pro Thr Ile Ile Asp Glu Asn Gly Val Lys Lys
    290                 295                 300

Glu Glu Glu Leu Met Val Glu Arg Thr Val Lys Leu Asn Arg Gln Gly
305                 310                 315                 320

Glu Ala Ala Val Val Phe Ser Asn Asp Glu Leu Lys Arg His Lys Leu
                325                 330                 335

Leu His Glu Trp Gly Gly Gly Ser Ile Arg Ile Val Ala Ser Val Thr
            340                 345                 350

Glu Asp Ile Thr Glu Ile Glu Arg Asn Ala Thr His Gln Ile Ser Thr
        355                 360                 365

Phe Arg Glu Glu Val Lys Leu Asp Val Glu Lys Gln Gly Asp Thr Phe
    370                 375                 380

Lys Pro Gly Leu Thr Tyr Asn Val Val Val Ala Leu Lys Gln Met Asp
385                 390                 395                 400

Asp Thr Pro Val Lys Ala Thr Leu Pro Lys Arg Val Gln Val Ser Thr
                405                 410                 415

Phe Tyr Asn Tyr Pro Tyr Asn His Asp Thr Ser Ser Leu Gln Glu Glu
            420                 425                 430

Lys Glu Thr Lys Ile Val Glu Val Asp Ala His Gly Thr Ser Val Leu
        435                 440                 445
```

```
Thr Leu Gln Pro Pro Ile Asn Cys Thr Ser Ala Arg Ile Glu Ala His
    450                 455                 460

Tyr Asp Ile Gly Gly Lys Asp Asn Phe Thr Ala Thr Pro Ile Tyr Ser
465                 470                 475                 480

Ser Leu Tyr Val Glu Ala Ala Val Ser Pro Thr Lys Ser Phe Leu Gln
                485                 490                 495

Leu Leu Ala Asp Asn Glu Gly Ala Val Asp Val Gly Lys Ser Leu Ser
            500                 505                 510

Phe Ser Leu Lys Ala Thr Gln Pro Leu Ser Thr Ile Thr Tyr Gln Val
        515                 520                 525

Met Ser Arg Ser Asn Ile Val Val Ser Gln Gln Met Thr Val Asn Ser
    530                 535                 540

Glu His Ala Thr Ile Ser Phe Pro Ala Thr Ala Asn Met Ala Pro Lys
545                 550                 555                 560

Ser Arg Leu Ile Val Tyr Ala Ile Ile Glu Ser Ser Gln Glu Val Leu
                565                 570                 575

Val Asp Ala Leu Asp Phe Lys Val Glu Gly Ile Phe Gln Asn Gln Val
            580                 585                 590

Ala Leu Ser Ile Asp Lys Gln Ala Val Glu Pro Gly Gln Asn Val Lys
        595                 600                 605

Phe Lys Val Thr Ser Asp Lys Asn Ser Phe Val Gly Leu Leu Val Val
    610                 615                 620

Asp Gln Ser Val Leu Leu Leu Lys Thr Gly Asn Asp Ile Thr Arg Glu
625                 630                 635                 640

Lys Val Glu Gln Asp Leu Glu Asn Tyr Asp Ser Asn Asn Val Gly Gly
                645                 650                 655

Gly Phe Gly Gly Pro Arg Pro Trp Glu Ala Ile Asp Arg Lys Lys Arg
            660                 665                 670

Ser Ile Trp Arg Pro Trp Trp Gly Ile Gly Gly Ser Asp Ala Gln Ser
        675                 680                 685

Ile Phe Ser Asn Ala Gly Leu Val Val Leu Thr Asp Ala Leu Leu Tyr
    690                 695                 700

Arg Glu Pro Gln Arg Glu Phe Met Ser Val Met Met Met Asp Gly Ala
705                 710                 715                 720

Pro Gly Met Ala Glu Ala Ala Phe Ala Ala Pro Pro Met Gly Gly Ser
                725                 730                 735

Ser Pro Pro Pro Pro Thr Val Arg Lys Phe Phe Pro His Thr Trp Ile
            740                 745                 750

Trp Ser Asp Leu Asn Ser Thr Ser Gly Glu Val Glu Met Glu Ile Glu
```

755 760 765

Ala Pro Asp Thr Ile Thr Ser Trp Val Ala Ser Thr Phe Ala Ile Asn
770 775 780

Glu Glu Asn Gly Leu Gly Val Ala Pro Thr Thr Ser Lys Leu Arg Val
785 790 795 800

Phe Arg Pro Phe Phe Ile Gln Leu Asn Leu Pro Tyr Ala Val Arg Arg
805 810 815

Gly Glu Lys Phe Ala Leu Leu Val Leu Val Phe Asn Tyr Met Glu Lys
820 825 830

Glu Gln Asp Val Thr Val Thr Leu Lys Tyr Asp Lys Asp Ser Gly Tyr
835 840 845

Asp Leu Leu Lys Lys Asp Gly Thr Val Val Arg Arg Asp Glu Val Gly
850 855 860

Gln Gln Asn Val Arg Ile Val Ser Val Ala Gly Gly Gly Thr Ser Lys
865 870 875 880

Ala Val Tyr Phe Pro Ile Val Pro Ser Ser Ile Gly Glu Ile Pro Val
885 890 895

His Ile Ser Ala Ile Ala Ser Gln Gly Gly Asp Ala Val Glu Met Asn
900 905 910

Leu Arg Val Asp Pro Gln Gly Tyr Lys Val Asp Arg Asn Ile Pro Phe
915 920 925

Val Ile Asp Leu Asn Asn Asn Ser Ser Asp Phe Ser Lys Asn Leu Glu
930 935 940

Leu Ile Trp Pro Asn Asp Val Val Asp Gly Ser Gln Lys Ala Arg Leu
945 950 955 960

Asp Val Ile Gly Asp Met Met Gly Pro Val Leu Asn Asn Ala His Lys
965 970 975

Leu Val Gln Met Pro Tyr Gly Cys Gly Glu Gln Asn Met Leu Asn Leu
980 985 990

Val Pro Asn Ile Leu Val Val Lys Tyr Leu Arg Ala Thr Asn Arg Asn
995 1000 1005

Glu Ser Gln Leu Glu Thr Lys Ala Ile Lys Phe Ile Glu Gln Gly Ile
1010 1015 1020

Gln Arg Glu Leu Thr Tyr Lys Arg Ala Asp Asn Ser Phe Ser Ala Phe
1025 1030 1035 1040

Gly Asp Ser Asp Lys Ala Gly Ser Thr Trp Leu Thr Ala Phe Val Val
1045 1050 1055

Arg Ser Phe His His Ala Lys Gln Tyr Ala Phe Val Asp Pro Asn Val
1060 1065 1070

Ile Ser Arg Ala Val Ala Phe Leu Asn Ser Gln Gln Met Glu Ser Gly
1075 1080 1085

Ala Phe Ala Glu Arg Gly Glu Val His His Lys Asp Met Gln Gly Gly
1090 1095 1100

Ala Gln Asp Gly Gly Val Ala Leu Thr Ala Phe Val Leu Ile Ser Ile
1105 1110 1115 1120

Leu Glu Asn Gly Met Glu Asn Gly Lys Ala Val Thr Tyr Leu Glu Lys
1125 1130 1135

His Leu Asp Glu Val Ser Gly Asn Ala Tyr Thr Met Ala Val Val Ala
1140 1145 1150

Tyr Ala Leu Gln Leu Ala Lys Ser Lys Gln Ala Gly Lys Ala Phe Glu
1155 1160 1165

Asn Leu Lys Lys His Lys Ile Val Glu Lys Ser Gly Asp Val Lys Phe
1170 1175 1180

Ala Ser Ala Gln Lys Lys Val Glu Lys Leu Lys Glu Ser Arg Ala Tyr
1185 1190 1195 1200

Met Phe Gln Ala Arg Pro Val Asp Ile Glu Thr Thr Ser Tyr Ala Val
1205 1210 1215

Leu Ser Tyr Leu Ala Gln Asn Gln Thr Ser Glu Ser Leu Ser Ile Ile
1220 1225 1230

Arg Trp Leu Val Ser Gln Arg Asn Glu Leu Gly Gly Phe Thr Ser Thr
1235 1240 1245

Gln Asp Thr Val Met Ala Leu Gln Ala Leu Ser Ser Tyr Ala Ala Val
1250 1255 1260

Thr Tyr Ser Asp Lys His Thr Ser Gln Val Thr Ile Leu Asn Gly Lys
1265 1270 1275 1280

His Thr His Ser Phe Asp Ile Asn Ile Arg Asn Ala Ile Val Leu Gln
1285 1290 1295

Ser Tyr Gln Leu Ser Ser Leu Asn Asp Ala Val Ser Ile Asn Ala Asn
1300 1305 1310

Gly Thr Gly Val Val Phe Ala Gln Leu Ser Tyr Ser Tyr Tyr Arg Asp
1315 1320 1325

Ser Leu Asn Asp Asp Ala Pro Phe Phe Cys Ser Gln Glu Ile Lys Glu
1330 1335 1340

Ile Arg Ala Gly Asn Arg Leu Gln Leu Asp Leu Cys Cys Asn Tyr Thr
1345 1350 1355 1360

Arg Pro Gly Lys Ser Asn Met Ala Leu Ala Glu Ile Asp Ala Leu Ser
1365 1370 1375

```
Gly Tyr Arg Phe Asp Ala Glu Gln Val His Thr Leu Thr Ser Ile Glu
            1380                1385                1390

Asp Leu Gln Arg Val Glu Met Glu Lys Asp Asp Thr Lys Met Asn Val
        1395                1400                1405

Tyr Phe Asn Pro Leu Gly Gly Arg Pro Val Cys Leu Ser Leu Tyr Ser
    1410                1415                1420

Asp Val Thr Tyr Gln Val Ala Asp Gln Lys Pro Ala Asn Phe Arg Leu
1425                1430                1435                1440

Val Asp Tyr Tyr Asp Pro Glu Glu Gln Leu Lys Met Thr Tyr Ala Ala
                1445                1450                1455

Lys Gln Thr Arg Ser Leu Gln Glu Lys Cys Gly Glu Asp Cys Trp Pro
            1460                1465                1470

Pro Ile Ser Pro Ser Leu Pro Pro Phe Asp Glu Ser Thr Val Thr Gly
        1475                1480                1485

Thr Ser Ser Gly Phe Gly Ala Lys Trp Cys Ala Leu Ile Ile Ala Val
    1490                1495                1500

Leu Leu Ile Ala
1505


<210> 71
<211> 1519
<212> PRT
<213> Caenorhabditis elegans

<400> 71
Met Arg Leu Leu Ile Leu Asn Ile Leu Phe Val Val Trp Gln Ile His
  1               5                  10                  15

Gly Val Ile Gly Gln Ser Thr Asn Ala Ala Val Val Ser Thr Thr Ala
             20                  25                  30

Ala Pro Val Lys Pro Ala Thr Tyr Met Leu Val Ala Pro Ala Val Val
         35                  40                  45

Arg Pro Asp Gln Pro Phe Ser Val Cys Met Asn Leu Leu Lys Gln Ala
     50                  55                  60

Thr Asp Glu Asp Met Ile Val Arg Ile Glu Val Arg Thr Glu Arg Asn
 65                  70                  75                  80

Glu Thr Ile Ala Ala Arg Val Ile Ser Asn Leu Lys Pro Gly Ile Ala
                 85                  90                  95

Gln Thr Val Ser Leu Ser Glu Met Pro Ala Gln Ser Leu Thr Pro Arg
            100                 105                 110

Gln Ser Tyr Lys Leu Tyr Ile Arg Gly Glu Thr Leu Asn Ala Glu Leu
        115                 120                 125
```

Ile Phe Glu Asn Glu Asn Glu Leu Lys Tyr Asp Gln Lys Ala Leu Ser
130 135 140

Val Phe Ile Gln Thr Asp Arg Ala Ile Tyr Arg Pro Ala Ser Leu Val
145 150 155 160

Arg Tyr Arg Ala Ile Val Val Lys Ser Asp Leu Lys Pro Tyr Val Gly
165 170 175

Asn Ala Thr Ile Lys Ile Phe Asp Pro Ser Arg Asn Leu Ile Ser Gln
180 185 190

Thr Ile Gly Val Thr Leu Asp Arg Gly Val Tyr Ser Gly Glu Leu Gln
195 200 205

Leu Ala Glu Glu Thr Leu Leu Gly Asp Trp Phe Ile Glu Val Glu Thr
210 215 220

Ser Asn Gly Val Gln Asp Lys Ser Ser Phe Thr Val Asp Thr Tyr Val
225 230 235 240

Leu Pro Lys Phe Glu Val Asn Ile Lys Thr Ser Ser Phe Ile Thr Ile
245 250 255

Asn Asp Asp Leu Ser Val Phe Val Asp Ala Lys Tyr Thr Tyr Gly Lys
260 265 270

Gly Val Ala Gly Lys Ala Lys Val Ser Leu Glu Leu Pro Trp His Arg
275 280 285

Trp His Ala Met Val Pro Thr Ile Ile Asp Glu Asn Gly Val Lys Lys
290 295 300

Glu Glu Glu Leu Met Val Glu Arg Thr Val Lys Leu Asn Arg Gln Gly
305 310 315 320

Glu Ala Ala Val Val Phe Ser Asn Asp Glu Leu Lys Arg His Lys Leu
325 330 335

Leu His Glu Trp Gly Gly Gly Ser Ile Arg Ile Val Ala Ser Val Thr
340 345 350

Glu Asp Ile Thr Glu Ile Glu Arg Asn Ala Thr His Gln Ile Ser Thr
355 360 365

Phe Arg Glu Glu Val Lys Leu Asp Val Glu Lys Gln Gly Asp Thr Phe
370 375 380

Lys Pro Gly Leu Thr Tyr Asn Val Val Val Ala Leu Lys Gln Met Asp
385 390 395 400

Asp Thr Pro Val Lys Ala Thr Leu Pro Lys Arg Val Gln Val Ser Thr
405 410 415

Phe Tyr Asn Tyr Pro Tyr Asn His Asp Thr Ser Ser Leu Gln Glu Glu
420 425 430

Lys Glu Thr Lys Ile Val Glu Val Asp Ala His Gly Thr Ser Val Leu

```
        435                 440                 445

Thr Leu Gln Pro Pro Ile Asn Cys Thr Ser Ala Arg Ile Glu Ala His
    450                 455                 460

Tyr Asp Ile Gly Gly Lys Asp Asn Phe Thr Ala Thr Pro Ile Tyr Ser
465                 470                 475                 480

Ser Leu Tyr Val Glu Ala Ala Val Ser Pro Thr Lys Ser Phe Leu Gln
                485                 490                 495

Leu Leu Ala Asp Asn Glu Gly Ala Val Asp Val Gly Lys Ser Leu Ser
            500                 505                 510

Phe Ser Leu Lys Ala Thr Gln Pro Leu Ser Thr Ile Thr Tyr Gln Val
        515                 520                 525

Met Ser Arg Ser Asn Ile Val Val Ser Gln Gln Met Thr Val Asn Ser
    530                 535                 540

Glu His Ala Thr Ile Ser Phe Pro Ala Thr Ala Asn Met Ala Pro Lys
545                 550                 555                 560

Ser Arg Leu Ile Val Tyr Ala Ile Ile Glu Ser Ser Gln Glu Val Leu
                565                 570                 575

Val Asp Ala Leu Asp Phe Lys Val Glu Gly Ile Phe Gln Asn Gln Val
            580                 585                 590

Ala Leu Ser Ile Asp Lys Gln Ala Val Glu Pro Gly Gln Asn Val Lys
        595                 600                 605

Phe Lys Val Thr Ser Asp Lys Asn Ser Phe Val Gly Leu Leu Val Val
    610                 615                 620

Asp Gln Ser Val Leu Leu Leu Lys Thr Gly Asn Asp Ile Thr Arg Glu
625                 630                 635                 640

Lys Val Glu Gln Asp Leu Glu Asn Tyr Asp Ser Asn Asn Val Gly Gly
                645                 650                 655

Gly Phe Gly Gly Pro Arg Pro Trp Glu Ala Ile Asp Arg Lys Lys Arg
            660                 665                 670

Ser Ile Trp Arg Pro Trp Trp Gly Ile Gly Gly Ser Asp Ala Gln Ser
        675                 680                 685

Ile Phe Ser Asn Ala Gly Leu Val Val Leu Thr Asp Ala Leu Leu Tyr
    690                 695                 700

Arg Glu Pro Gln Arg Glu Phe Met Ser Glu Arg Arg Leu Asn Thr Pro
705                 710                 715                 720

Gly Gly Leu Thr Val Met Met Met Asp Gly Ala Pro Gly Met Ala Glu
                725                 730                 735

Ala Ala Phe Ala Ala Pro Pro Met Gly Gly Ser Ser Pro Pro Pro Pro
            740                 745                 750
```

Thr Val Arg Lys Phe Phe Pro His Thr Trp Ile Trp Ser Asp Leu Asn
755 760 765

Ser Thr Ser Gly Glu Val Glu Met Glu Ile Glu Ala Pro Asp Thr Ile
770 775 780

Thr Ser Trp Val Ala Ser Thr Phe Ala Ile Asn Glu Glu Asn Gly Leu
785 790 795 800

Gly Val Ala Pro Thr Thr Ser Lys Leu Arg Val Phe Arg Pro Phe Phe
805 810 815

Ile Gln Leu Asn Leu Pro Tyr Ala Val Arg Arg Gly Glu Lys Phe Ala
820 825 830

Leu Leu Val Leu Val Phe Asn Tyr Met Glu Lys Glu Gln Asp Val Thr
835 840 845

Val Thr Leu Lys Tyr Asp Lys Asp Ser Gly Tyr Asp Leu Leu Lys Lys
850 855 860

Asp Gly Thr Val Val Arg Arg Asp Glu Val Gly Gln Gln Asn Val Arg
865 870 875 880

Ile Val Ser Val Ala Gly Gly Gly Thr Ser Lys Ala Val Tyr Phe Pro
885 890 895

Ile Val Pro Ser Ser Ile Gly Glu Ile Pro Val His Ile Ser Ala Ile
900 905 910

Ala Ser Gln Gly Gly Asp Ala Val Glu Met Asn Leu Arg Val Asp Pro
915 920 925

Gln Gly Tyr Lys Val Asp Arg Asn Ile Pro Phe Val Ile Asp Leu Asn
930 935 940

Asn Asn Ser Ser Asp Phe Ser Lys Asn Leu Glu Leu Ile Trp Pro Asn
945 950 955 960

Asp Val Val Asp Gly Ser Gln Lys Ala Arg Leu Asp Val Ile Gly Asp
965 970 975

Met Met Gly Pro Val Leu Asn Asn Ala His Lys Leu Val Gln Met Pro
980 985 990

Tyr Gly Cys Gly Glu Gln Asn Met Leu Asn Leu Val Pro Asn Ile Leu
995 1000 1005

Val Val Lys Tyr Leu Arg Ala Thr Asn Arg Asn Glu Ser Gln Leu Glu
1010 1015 1020

Thr Lys Ala Ile Lys Phe Ile Glu Gln Gly Ile Gln Arg Glu Leu Thr
1025 1030 1035 1040

Tyr Lys Arg Ala Asp Asn Ser Phe Ser Ala Phe Gly Asp Ser Asp Lys
1045 1050 1055

Ala Gly Ser Thr Trp Leu Thr Ala Phe Val Val Arg Ser Phe His His
1060 1065 1070

Ala Lys Gln Tyr Ala Phe Val Asp Pro Asn Val Ile Ser Arg Ala Val
1075 1080 1085

Ala Phe Leu Asn Ser Gln Gln Met Glu Ser Gly Ala Phe Ala Glu Arg
1090 1095 1100

Gly Glu Val His His Lys Asp Met Gln Gly Gly Ala Gln Asp Gly Gly
1105 1110 1115 1120

Val Ala Leu Thr Ala Phe Val Leu Ile Ser Ile Leu Glu Asn Gly Met
1125 1130 1135

Glu Asn Gly Lys Ala Val Thr Tyr Leu Glu Lys His Leu Asp Glu Val
1140 1145 1150

Ser Gly Asn Ala Tyr Thr Met Ala Val Val Ala Tyr Ala Leu Gln Leu
1155 1160 1165

Ala Lys Ser Lys Gln Ala Gly Lys Ala Phe Glu Asn Leu Lys Lys His
1170 1175 1180

Lys Ile Val Glu Lys Ser Gly Asp Val Lys Phe Ala Ser Ala Gln Lys
1185 1190 1195 1200

Lys Val Glu Lys Leu Lys Glu Ser Arg Ala Tyr Met Phe Gln Ala Arg
1205 1210 1215

Pro Val Asp Ile Glu Thr Thr Ser Tyr Ala Val Leu Ser Tyr Leu Ala
1220 1225 1230

Gln Asn Gln Thr Ser Glu Ser Leu Ser Ile Ile Arg Trp Leu Val Ser
1235 1240 1245

Gln Arg Asn Glu Leu Gly Gly Phe Thr Ser Thr Gln Asp Thr Val Met
1250 1255 1260

Ala Leu Gln Ala Leu Ser Ser Tyr Ala Ala Val Thr Tyr Ser Asp Lys
1265 1270 1275 1280

His Thr Ser Gln Val Thr Ile Leu Asn Gly Lys His Thr His Ser Phe
1285 1290 1295

Asp Ile Asn Ile Arg Asn Ala Ile Val Leu Gln Ser Tyr Gln Leu Ser
1300 1305 1310

Ser Leu Asn Asp Ala Val Ser Ile Asn Ala Asn Gly Thr Gly Val Val
1315 1320 1325

Phe Ala Gln Leu Ser Tyr Ser Tyr Tyr Arg Asp Ser Leu Asn Asp Asp
1330 1335 1340

Ala Pro Phe Phe Cys Ser Gln Glu Ile Lys Glu Ile Arg Ala Gly Asn
1345 1350 1355 1360

Arg Leu Gln Leu Asp Leu Cys Cys Asn Tyr Thr Arg Pro Gly Lys Ser

```
                1365                    1370                    1375

Asn Met Ala Leu Ala Glu Ile Asp Ala Leu Ser Gly Tyr Arg Phe Asp
            1380                    1385                    1390

Ala Glu Gln Val His Thr Leu Thr Ser Ile Glu Asp Leu Gln Arg Val
        1395                    1400                    1405

Glu Met Glu Lys Asp Asp Thr Lys Met Asn Val Tyr Phe Asn Pro Leu
    1410                    1415                    1420

Gly Gly Arg Pro Val Cys Leu Ser Leu Tyr Ser Asp Val Thr Tyr Gln
1425                    1430                    1435                    1440

Val Ala Asp Gln Lys Pro Ala Asn Phe Arg Leu Val Asp Tyr Tyr Asp
                1445                    1450                    1455

Pro Glu Glu Gln Leu Lys Met Thr Tyr Ala Ala Lys Gln Thr Arg Ser
            1460                    1465                    1470

Leu Gln Glu Lys Cys Gly Glu Asp Cys Trp Pro Pro Ile Ser Pro Ser
        1475                    1480                    1485

Leu Pro Pro Phe Asp Glu Ser Thr Val Thr Gly Thr Ser Ser Gly Phe
    1490                    1495                    1500

Gly Ala Lys Trp Cys Ala Leu Ile Ile Ala Val Leu Leu Ile Ala
1505                    1510                    1515
```

<210> 72
<211> 1026
<212> DNA
<213> Caenorhabditis elegans

<400> 72

```
atgaatattg tgtgggttat aatttttgg aagctacaga aaggcatttt ccgagaagat 60
ggacttgaac cagtaacact agcagttcat ggtactgcag aaatgatgga aatgatagaa 120
aataagcctg agaactggga tgggcctata tcatttggat tgtttattga ttttcattct 180
agacaaattc tggattacgt cgctaaggtt tatagctgcg atgaggagtt tcagaaaaag 240
gttaccgtac actttgcatt ccgtctatca ccctttcaaa ctagctgccc acaaatcaaa 300
gtctcaccgt caactctcga atgtggcgag ttcctctcca acagaaaaaa gtttcgacgt 360
gctgtaggcg actcatttca attgtaccca agtaacttga tgagaaatat tgcaagaaaa 420
ggtgccaaat cggatattca tttcattgtc gatggtgata tgataatgag tgatggattt 480
gcggaaaaga ttaaaccaat tgcaaatcaa attgttgatg ggaaaaacaa aaatgtattg 540
gttgttagaa gatttgaaac aaatgaaacg acaatacccc acaaccatat agaattgaaa 600
aatgcaattg aaaataaaca agtcttccaa tttcatcaca gatttttctt tgctgggcat 660
aaaatttcaa atatttcaca ctggtttgcg gtttccaatg aaaccgacga aattactgct 720
tgggaaatac cttactcgag cagcttatgg gaagttcaag tgattcttca tcgtaacgac 780
ttgtacaatg cggattactt tccagcgaga atcaaggtta tgcagtcatt ggtgtacagc 840
ctgtgcagag ccaactatac gttcaacctc ctatcacacg ttttcaacgt gcacaaagga 900
atcaaattag gcgatactaa cttctcaaaa tccgttatag cacactctaa gagaaatgga 960
agaaatagcg aactccagga tacatacca gatacgttgg atagatgcgg tcaatttgta 1020
atgtaa                                                            1026
```

<210> 73
<211> 341
<212> PRT

<213> Caenorhabditis elegans

<400> 73

Met Asn Ile Val Trp Val Ile Ile Phe Trp Lys Leu Gln Lys Gly Ile
1 5 10 15

Phe Arg Glu Asp Gly Leu Glu Pro Val Thr Leu Ala Val His Gly Thr
20 25 30

Ala Glu Met Met Glu Met Ile Glu Asn Lys Pro Glu Asn Trp Asp Gly
35 40 45

Pro Ile Ser Phe Gly Leu Phe Ile Asp Phe His Ser Arg Gln Ile Leu
50 55 60

Asp Tyr Val Ala Lys Val Tyr Ser Cys Asp Glu Glu Phe Gln Lys Lys
65 70 75 80

Val Thr Val His Phe Ala Phe Arg Leu Ser Pro Phe Gln Thr Ser Cys
85 90 95

Pro Gln Ile Lys Val Ser Pro Ser Thr Leu Glu Cys Gly Glu Phe Leu
100 105 110

Ser Asn Arg Lys Lys Phe Arg Arg Ala Val Gly Asp Ser Phe Gln Leu
115 120 125

Tyr Pro Ser Asn Leu Met Arg Asn Ile Ala Arg Lys Gly Ala Lys Ser
130 135 140

Asp Ile His Phe Ile Val Asp Gly Asp Met Ile Met Ser Asp Gly Phe
145 150 155 160

Ala Glu Lys Ile Lys Pro Ile Ala Asn Gln Ile Val Asp Gly Lys Asn
165 170 175

Lys Asn Val Leu Val Val Arg Arg Phe Glu Thr Asn Glu Thr Thr Ile
180 185 190

Pro His Asn His Ile Glu Leu Lys Asn Ala Ile Glu Asn Lys Gln Val
195 200 205

Phe Gln Phe His His Arg Phe Phe Phe Ala Gly His Lys Ile Ser Asn
210 215 220

Ile Ser His Trp Phe Ala Val Ser Asn Glu Thr Asp Glu Ile Thr Ala
225 230 235 240

Trp Glu Ile Pro Tyr Ser Ser Ser Leu Trp Glu Val Gln Val Ile Leu
245 250 255

His Arg Asn Asp Leu Tyr Asn Ala Asp Tyr Phe Pro Ala Arg Ile Lys
260 265 270

Val Met Gln Ser Leu Val Tyr Ser Leu Cys Arg Ala Asn Tyr Thr Phe
275 280 285

Asn Leu Leu Ser His Val Phe Asn Val His Lys Gly Ile Lys Leu Gly

290 295 300

Asp Thr Asn Phe Ser Lys Ser Val Ile Ala His Ser Lys Arg Asn Gly
305 310 315 320

Arg Asn Ser Glu Leu Gln Asp Thr Tyr Pro Asp Thr Leu Asp Arg Cys
325 330 335

Gly Gln Phe Val Met
340

<210> 74
<211> 1869
<212> DNA
<213> Caenorhabditis elegans

<400> 74
atgcagtaca ttgtagcgtc atattttaca atatggaact ttgtggacca cacacgtgtt 60
ggggcgtttc cggaagaaga ttacattcgc cttgcgtata ttatcggtgg aaattttatg 120
acacgtctta tgttcatgca gcattttaaa agtgtattga agtactctga ccatttcttt 180
cgcctacatc tgataactga tgaaaatcac cgatccgata ttcatgagct catgacgtca 240
tggaatattt caaactgcga gtggtttttt cacaatctca cagagttcga aaaacgagtt 300
gcttggattc caaattcaca ttattcaaaa tattatggtt tgagcaaact tttaattcct 360
gaaatcatcg gcaatgacat tggaaagatt atgttcatgg atgttgatat cattttttcaa 420
accaatattt ttgatttgtg gaaacaattt agaaacttta acaattccca ggttttcgga 480
atggtggaaa atctttccga ttggtatctt aacaaggatg gtaaaaagtc ggtttggcca 540
gcattgggac gtggatttaa cactggaata atcatgtttg atttggataa gctccggaaa 600
aatggatggg caagcaaatg gagagtagtg gctaacaaat atttacgaat tcacggaaaa 660
actgccatgt cagatcaaga catatttaat gcatacattc atgactatcc gactgaaatt 720
atacaaattc catgtgctta caattatcaa cttggagcac tcacaaaatc caaagaacta 780
tgtccggaga ctccacttgc gctacatttc aactcacaaa acaagactgt tggaaaaaat 840
tacgcatttt ttgacaagat tcgaaaagcg tttgatgaaa tggatggttc agatcttaaa 900
cgacgtagac gatcttttaa aggaaataat caaaaagata tctgccatga atatttacca 960
ttagacaatt tccgaattat cccaaatgca attggaagaa tgacaaaacc tgcagaactc 1020
tgcatggtga cgcagttttc caaagatcga cttaatcatt ttctggaaag tgcaaatgcg 1080
tggcggcatc ctatatctac agcagtatac gggaaggata aagatttgct agacattgca 1140
aaagccgtca cagaattgaa tcgcacagat atcacaatcc atctggtttt tgaggaacca 1200
actgaatcct ggatgctaga ctctttatat cccataaatt ttttgagaaa tgtagcaatt 1260
gaacacgcaa actgcaaata tattttaatg actgatgtag attttgtggt acttggagat 1320
tacggtacaa taatcgatca aactggaaac ctgaaacaaa aagaagttct ggtgatccct 1380
gcattggaaa tgacataccc gcagctgagg ttaaatttat caaattttct ctcgagaaaa 1440
gatttggtga tagaacattt attgaacaaa accattcaaa cttttcggga aacaatttgg 1500
ccaagttctc acgtacctac aaatatttca aaatggataa agtcaaatcg cacttatatg 1560
gtagcacaaa atgtaaatta tgaaaagaac tatgaacctt attttgtaat caaaaaagaa 1620
gaatgccctt tttacgatca gcggttcgga ggatttggct ggaataaagt aactcatgta 1680
atgcaactga aaatgatgaa ttataagttc ttggtttctc caactagttt tatgattcat 1740
caaaatcata atgcttcaaa atcattgaaa cgctggagac gtgacccaca ttatcagaaa 1800
tgccttcaca cgttaaaaaa taaatttatg aaaaaaactg ctagtcgact gggaatcaag 1860
ctaagatga 1869

<210> 75
<211> 622
<212> PRT
<213> Caenorhabditis elegans

<400> 75
Met Gln Tyr Ile Val Ala Ser Tyr Phe Thr Ile Trp Asn Phe Val Asp

1 5 10 15

His Thr Arg Val Gly Ala Phe Pro Glu Glu Asp Tyr Ile Arg Leu Ala
20 25 30

Tyr Ile Ile Gly Gly Asn Phe Met Thr Arg Leu Met Phe Met Gln His
35 40 45

Phe Lys Ser Val Leu Lys Tyr Ser Asp His Phe Phe Arg Leu His Leu
50 55 60

Ile Thr Asp Glu Asn His Arg Ser Asp Ile His Glu Leu Met Thr Ser
65 70 75 80

Trp Asn Ile Ser Asn Cys Glu Trp Phe Phe His Asn Leu Thr Glu Phe
85 90 95

Glu Lys Arg Val Ala Trp Ile Pro Asn Ser His Tyr Ser Lys Tyr Tyr
100 105 110

Gly Leu Ser Lys Leu Leu Ile Pro Glu Ile Ile Gly Asn Asp Ile Gly
115 120 125

Lys Ile Met Phe Met Asp Val Asp Ile Ile Phe Gln Thr Asn Ile Phe
130 135 140

Asp Leu Trp Lys Gln Phe Arg Asn Phe Asn Asn Ser Gln Val Phe Gly
145 150 155 160

Met Val Glu Asn Leu Ser Asp Trp Tyr Leu Asn Lys Asp Gly Lys Lys
165 170 175

Ser Val Trp Pro Ala Leu Gly Arg Gly Phe Asn Thr Gly Ile Ile Met
180 185 190

Phe Asp Leu Asp Lys Leu Arg Lys Asn Gly Trp Ala Ser Lys Trp Arg
195 200 205

Val Val Ala Asn Lys Tyr Leu Arg Ile His Gly Lys Thr Ala Met Ser
210 215 220

Asp Gln Asp Ile Phe Asn Ala Tyr Ile His Asp Tyr Pro Thr Glu Ile
225 230 235 240

Ile Gln Ile Pro Cys Ala Tyr Asn Tyr Gln Leu Gly Ala Leu Thr Lys
245 250 255

Ser Lys Glu Leu Cys Pro Glu Thr Pro Leu Ala Leu His Phe Asn Ser
260 265 270

Gln Asn Lys Thr Val Gly Lys Asn Tyr Ala Phe Phe Asp Lys Ile Arg
275 280 285

Lys Ala Phe Asp Glu Met Asp Gly Ser Asp Leu Lys Arg Arg Arg Arg
290 295 300

Ser Phe Lys Gly Asn Asn Gln Lys Asp Ile Cys His Glu Tyr Leu Pro
305 310 315 320

Leu Asp Asn Phe Arg Ile Ile Pro Asn Ala Ile Gly Arg Met Thr Lys
325 330 335

Pro Ala Glu Leu Cys Met Val Thr Gln Phe Ser Lys Asp Arg Leu Asn
340 345 350

His Phe Leu Glu Ser Ala Asn Ala Trp Arg His Pro Ile Ser Thr Ala
355 360 365

Val Tyr Gly Lys Asp Lys Asp Leu Leu Asp Ile Ala Lys Ala Val Thr
370 375 380

Glu Leu Asn Arg Thr Asp Ile Thr Ile His Leu Val Phe Glu Glu Pro
385 390 395 400

Thr Glu Ser Trp Met Leu Asp Ser Leu Tyr Pro Ile Asn Phe Leu Arg
405 410 415

Asn Val Ala Ile Glu His Ala Asn Cys Lys Tyr Ile Leu Met Thr Asp
420 425 430

Val Asp Phe Val Val Leu Gly Asp Tyr Gly Thr Ile Ile Asp Gln Thr
435 440 445

Gly Asn Leu Lys Gln Lys Glu Val Leu Val Ile Pro Ala Leu Glu Met
450 455 460

Thr Tyr Pro Gln Leu Arg Leu Asn Leu Ser Asn Phe Leu Ser Arg Lys
465 470 475 480

Asp Leu Val Ile Glu His Leu Leu Asn Lys Thr Ile Gln Thr Phe Arg
485 490 495

Glu Thr Ile Trp Pro Ser Ser His Val Pro Thr Asn Ile Ser Lys Trp
500 505 510

Ile Lys Ser Asn Arg Thr Tyr Met Val Ala Gln Asn Val Asn Tyr Glu
515 520 525

Lys Asn Tyr Glu Pro Tyr Phe Val Ile Lys Lys Glu Glu Cys Pro Phe
530 535 540

Tyr Asp Gln Arg Phe Gly Gly Phe Gly Trp Asn Lys Val Thr His Val
545 550 555 560

Met Gln Leu Lys Met Met Asn Tyr Lys Phe Leu Val Ser Pro Thr Ser
565 570 575

Phe Met Ile His Gln Asn His Asn Ala Ser Lys Ser Leu Lys Arg Trp
580 585 590

Arg Arg Asp Pro His Tyr Gln Lys Cys Leu His Thr Leu Lys Asn Lys
595 600 605

Phe Met Lys Lys Thr Ala Ser Arg Leu Gly Ile Lys Leu Arg
610 615 620

<210> 76
<211> 417
<212> PRT
<213> Caenorhabditis elegans

<400> 76
Met Val Ser Leu Gln Lys Ser Ile Gly Leu Leu Leu Leu Ser Ala Ile
1               5                   10                  15

Ile Gly Leu Val Phe Leu Ile Gln His Arg Lys Ser Tyr Thr Ser Ser
            20                  25                  30

Asp Ala Leu Leu Glu Asn Gly Tyr Pro Asn Lys Tyr Tyr Thr Ile Glu
        35                  40                  45

Asn Pro Ala Glu Glu Gly Glu Arg Arg Ser Tyr Ser Ile Gln Thr Glu
    50                  55                  60

Met His Ala Asp Gln Tyr Cys Ile Ala Tyr Lys Phe Leu Glu Ala Thr
65                  70                  75                  80

Glu Ser Phe Arg Glu Ala Asp Gly Leu Glu Pro Val Thr Leu Ala Thr
                85                  90                  95

His Ala Thr Ala Asp Met Ile Glu Thr Val Glu Asn Met Thr Phe Leu
            100                 105                 110

Trp Asp Gly Pro Ile Ser Ile Gly Ile Phe Val Asp Tyr His Ser Tyr
        115                 120                 125

Asn Val Leu Glu Tyr Leu Ala Glu Val His Arg Cys Asp Val Ser Phe
    130                 135                 140

Arg Arg Lys Met Asn Val His Phe Ala Phe Arg Arg Ser Pro Phe Gln
145                 150                 155                 160

Thr Glu Cys Pro Leu Ile Glu Ile Pro Gln Ser Asn Arg Ser Cys Gln
                165                 170                 175

Glu Phe Phe Ala Thr His Thr Glu Leu Arg Asn Ala Ile Val Gly Pro
            180                 185                 190

Phe Gln Leu Tyr Pro Ser Asn Leu Met Arg Asn Ile Ala Arg Lys Gly
        195                 200                 205

Ala Gln Thr Asp Leu Gln Phe Ile Met Asp Gly Asp Met Val Pro Ser
    210                 215                 220

Glu Gly Phe Ala Thr Lys Ile Lys Arg Ile Ala Asn Glu Val Ile Asp
225                 230                 235                 240

Gly Lys Asn Lys Arg Val Leu Ala Ile Arg Arg Phe Glu Thr Ser Asp
                245                 250                 255

Thr Ala Glu Ile Pro Arg Asp His Leu Lys Leu Leu Lys Ser Lys Lys
            260                 265                 270

Leu His Lys Thr Phe Glu Phe His His Arg Tyr Phe Pro Glu Gly His
275 280 285

His Ile Asp Gly Leu Asp Asp Trp Phe Arg Thr Ser Ile His Ser Gly
290 295 300

Val Val Thr Thr Lys Glu Val Ala Tyr Pro Gly Tyr Leu Trp Glu Val
305 310 315 320

Gln Thr Ile Leu His Arg Asn Asp Pro Tyr Asn Ala Asp Tyr Phe Pro
325 330 335

Ser Arg Ile Lys Val Met His Ser Leu Val Tyr Ala Leu Cys Arg Ala
340 345 350

Gly Tyr Thr Phe His Val Pro Thr His Val Phe Asp Ser His Arg Gly
355 360 365

Ile Lys His Thr Asn Thr Ile Tyr Ser Lys Ala Thr Ile Ala His Gln
370 375 380

Glu Ala Tyr Ala Met Lys Glu Ala Gly Asp Arg Tyr Ile Lys Glu Met
385 390 395 400

Asp Asp Leu Tyr Pro His Thr Leu Ser Gln Cys Gly Glu Phe Ser Met
405 410 415

Ile

<210> 77
<211> 1050
<212> DNA
<213> Caenorhabditis elegans

<400> 77
atgcacgacg agcagttttg tgttggttat aacttcctgg aggccgaaga tacattccga 60
gaagacggct tggaaccagt aactcttgct attcatggca ctccagaagt tttgcaacta 120
ttggggaata agcctttgaa ctgggatggg cctatttcat ttggtctatt cgtagatttt 180
cattcccaaa aggccctgaa ttatatttcc atgctacata aatgtgatgc agcttttaaa 240
agacagatga ctgtccactt tgcattccga atctcaccat ctcaatccga atgcccaatg 300
attcaagttc ttgggtatca ggattgtgcc acatttttac agaaaagcaa gcagctcctt 360
gaggaaattg aggactcttt tcaaatctac ccgataaacc taatgagaaa tattgctcgg 420
cgcggagcaa agtcggattt acacttgata atcgatacag atatgatgat gagcaccaac 480
tttgcaaaga tggtaaaacc aatcgcaaat cggatgattg atgggaagaa taagcaagtg 540
ttggttgtta gacgtttcga gaccaacgaa aatgagctac caatgagctt tggggatctt 600
gaggagggaa ttgaaaatca taaaacattc cagttccatc acaaattctt tttcgttggt 660
catcaaattc ccaacttgat ggaatggttc gaaagatctc acgcctctaa tgatgtggtg 720
gcatgggaga ttccatacac tggaaatgat tgggaagttc aaatcattct tcaccgcaac 780
gatccatata atgtagagta ttttccttcg cgagtcaagg atatgcagtc tttgatctac 840
aagttatgcc gtgcaaacta caccttcaat ttgctctctc atgtgttcaa tgttcataaa 900
ggaatcaaag aagatgatac catgtactcg aaggttgtga ctgctcacac aaaacggcaa 960
ggaagattga ggacgctttc tcgatatgtc actgaaattg acaggaaata cccggatacc 1020
atgaaacgat gtgggcagtt tttgttataa 1050

<210> 78
<211> 349

<212> PRT
<213> Caenorhabditis elegans

<400> 78
Met His Asp Glu Gln Phe Cys Val Gly Tyr Asn Phe Leu Glu Ala Glu
1 5 10 15

Asp Thr Phe Arg Glu Asp Gly Leu Glu Pro Val Thr Leu Ala Ile His
20 25 30

Gly Thr Pro Glu Val Leu Gln Leu Leu Gly Asn Lys Pro Leu Asn Trp
35 40 45

Asp Gly Pro Ile Ser Phe Gly Leu Phe Val Asp Phe His Ser Gln Lys
50 55 60

Ala Leu Asn Tyr Ile Ser Met Leu His Lys Cys Asp Ala Ala Phe Lys
65 70 75 80

Arg Gln Met Thr Val His Phe Ala Phe Arg Ile Ser Pro Ser Gln Ser
85 90 95

Glu Cys Pro Met Ile Gln Val Leu Gly Tyr Gln Asp Cys Ala Thr Phe
100 105 110

Leu Gln Lys Ser Lys Gln Leu Leu Glu Glu Ile Glu Asp Ser Phe Gln
115 120 125

Ile Tyr Pro Ile Asn Leu Met Arg Asn Ile Ala Arg Arg Gly Ala Lys
130 135 140

Ser Asp Leu His Leu Ile Ile Asp Thr Asp Met Met Met Ser Thr Asn
145 150 155 160

Phe Ala Lys Met Val Lys Pro Ile Ala Asn Arg Met Ile Asp Gly Lys
165 170 175

Asn Lys Gln Val Leu Val Val Arg Arg Phe Glu Thr Asn Glu Asn Glu
180 185 190

Leu Pro Met Ser Phe Gly Asp Leu Glu Glu Gly Ile Glu Asn His Lys
195 200 205

Thr Phe Gln Phe His His Lys Phe Phe Phe Val Gly His Gln Ile Pro
210 215 220

Asn Leu Met Glu Trp Phe Glu Arg Ser His Ala Ser Asn Asp Val Val
225 230 235 240

Ala Trp Glu Ile Pro Tyr Thr Gly Asn Asp Trp Glu Val Gln Ile Ile
245 250 255

Leu His Arg Asn Asp Pro Tyr Asn Val Glu Tyr Phe Pro Ser Arg Val
260 265 270

Lys Asp Met Gln Ser Leu Ile Tyr Lys Leu Cys Arg Ala Asn Tyr Thr
275 280 285

Phe Asn Leu Leu Ser His Val Phe Asn Val His Lys Gly Ile Lys Glu
290 295 300

Asp Asp Thr Met Tyr Ser Lys Val Val Thr Ala His Thr Lys Arg Gln
305 310 315 320

Gly Arg Leu Arg Thr Leu Ser Arg Tyr Val Thr Glu Ile Asp Arg Lys
325 330 335

Tyr Pro Asp Thr Met Lys Arg Cys Gly Gln Phe Leu Leu
340 345

<210> 79
<211> 1167
<212> DNA
<213> Caenorhabditis elegans

<400> 79
atgctcaaga tttcctcaag atttactcca tttgctttgt ttctcctatt ttcaattcta 60
ctttgtttgt ggtttttgaa gaaatattct caagatcttt ctaggatctc tatagaactt 120
tatgaaaatg agttttgcat tggctacaat ttcctggagg ctacagaaaa attccgagaa 180
gacggcttgg agcctgtgac acttgccatt catgggacat ccgatgtcct tgaagtagtg 240
gagaagaagc catcaaactg ggatgggcct atatcattcg ggatgtttgt tgactatcac 300
tcccagaagg ctctggaata tgtggcaatg cttcatcagt gtgataagga gttcggggag 360
aaagtcaccg ttcactatgt gttccgaact tctccttccc agatggattg tccagtgata 420
actcctgatg tgtcggtgaa ttgtgatgaa tttcgtcgga atcgaaagca gctcctcaaa 480
caaataacct ccccgtttca aatctaccca ataaacttga tgagaaatgt tgcccgccgt 540
ggagcaactt ctgatctaca cttgatagtc gacgctgata tgacaatgag ctctgatttt 600
gcgagaaaag tgaagccaat cgcaaatcgc ataattgatg ggaaacagag acaagttttg 660
gtagttcgac gttttgagac aaacgaagat gagattccac tggaagttga gcagctgaag 720
atgggatttg agaatcaaaa agtattcgag ttccatcaca atttttctt tattgggcat 780
aaaattccag acgtggaaaa atggtttcac gcctcgaaga cagaaaatga agtgactgcc 840
tgggaaattc catactcagg aaatgcatgg gaagtgcaag tgattcttca ccggaacgac 900
atgtacaacg ccgagtactt tccgtctaga atccgagaca tgcagtcttt gatctacggt 960
ctctgccgag ccaactacac cttcaacttg ctctctcacg tattcaatgt tcaccaaggc 1020
atcaaagagg atgacacaat gtactcgaaa gttgtcacag ctcactcgaa gcgatatgga 1080
aggaatagag cattctcccg ctacgtccat gagatgaata ctgcgtatcc gggaactatt 1140
cagcggtgcg ggaagtttga gatgtga 1167

<210> 80
<211> 388
<212> PRT
<213> Caenorhabditis elegans

<400> 80
Met Leu Lys Ile Ser Ser Arg Phe Thr Pro Phe Ala Leu Phe Leu Leu
1 5 10 15

Phe Ser Ile Leu Leu Cys Leu Trp Phe Leu Lys Lys Tyr Ser Gln Asp
20 25 30

Leu Ser Arg Ile Ser Ile Glu Leu Tyr Glu Asn Glu Phe Cys Ile Gly
35 40 45

Tyr Asn Phe Leu Glu Ala Thr Glu Lys Phe Arg Glu Asp Gly Leu Glu
50 55 60

Pro Val Thr Leu Ala Ile His Gly Thr Ser Asp Val Leu Glu Val Val
65 70 75 80

Glu Lys Lys Pro Ser Asn Trp Asp Gly Pro Ile Ser Phe Gly Met Phe
85 90 95

Val Asp Tyr His Ser Gln Lys Ala Leu Glu Tyr Val Ala Met Leu His
100 105 110

Gln Cys Asp Lys Glu Phe Gly Glu Lys Val Thr Val His Tyr Val Phe
115 120 125

Arg Thr Ser Pro Ser Gln Met Asp Cys Pro Val Ile Thr Pro Asp Val
130 135 140

Ser Val Asn Cys Asp Glu Phe Arg Arg Asn Arg Lys Gln Leu Leu Lys
145 150 155 160

Glu Ile Thr Ser Pro Phe Gln Ile Tyr Pro Ile Asn Leu Met Arg Asn
165 170 175

Val Ala Arg Arg Gly Ala Thr Ser Asp Leu His Leu Ile Val Asp Ala
180 185 190

Asp Met Thr Met Ser Ser Asp Phe Ala Arg Lys Val Lys Pro Ile Ala
195 200 205

Asn Arg Ile Ile Asp Gly Lys Gln Arg Gln Val Leu Val Val Arg Arg
210 215 220

Phe Glu Thr Asn Glu Asp Glu Ile Pro Leu Glu Val Glu Gln Leu Lys
225 230 235 240

Met Gly Phe Glu Asn Gln Lys Val Phe Glu Phe His His Asn Phe Phe
245 250 255

Phe Ile Gly His Lys Ile Pro Asp Val Glu Lys Trp Phe His Ala Ser
260 265 270

Lys Thr Glu Asn Glu Val Thr Ala Trp Glu Ile Pro Tyr Ser Gly Asn
275 280 285

Ala Trp Glu Val Gln Val Ile Leu His Arg Asn Asp Met Tyr Asn Ala
290 295 300

Glu Tyr Phe Pro Ser Arg Ile Arg Asp Met Gln Ser Leu Ile Tyr Gly
305 310 315 320

Leu Cys Arg Ala Asn Tyr Thr Phe Asn Leu Leu Ser His Val Phe Asn
325 330 335

Val His Gln Gly Ile Lys Glu Asp Asp Thr Met Tyr Ser Lys Val Val
340 345 350

Thr Ala His Ser Lys Arg Tyr Gly Arg Asn Arg Ala Phe Ser Arg Tyr
355 360 365

Val His Glu Met Asn Thr Ala Tyr Pro Gly Thr Ile Gln Arg Cys Gly

370 375 380

Lys Phe Glu Met
385

<210> 81
<211> 1275
<212> DNA
<213> Caenorhabditis elegans

<400> 81
atgtgtactt ttaaaaaatt cgatggagaa actcgaaaaa ctcggatcca gattttgtat 60
tttgctgcat cactcgttaa tctcgatctc aagcctgtga agctaaattc caatgcgaat 120
atatgcgtca aaattgaaac ttctcacttt acctctggca cttattatat taacttggca 180
tctgtacaat tcaaaggtaa tgctcctggt tctgatgctg aaggaaggtt tttcaagaaa 240
ctacacggaa aaccagaaaa taattataat tccttacaga cgactgtaca ctttgcattc 300
cgaatctcac catctcaaac cgaatgtcct gtgatctata cttccgggta taaggattgt 360
gtcacgtttt tccaaaagaa cacagagctc cttgaggaaa tggaggaccc ttttcagatc 420
tacccgataa atctaatgag aaatattgct cgacgcggag caaagtcgga tttacacttg 480
atagtcgata cagatatggt aatgagtact aactttgcaa agatggtaaa accagttgcg 540
aatcggatga ttgatgggat gaataaacaa gtcttggttg ttcgacgctt cgagaccaac 600
gaaaccgaac ttccactgaa cttggacgaa cttgagcaag ggcttctgaa tgagaacaca 660
tttgaattcc atcactcgtt ctttttttgtt ggccatcaaa tacccaactt gtctgagtgg 720
tttgaaaatt cttacgcatc agaagaaacc actgcatggg agattccata cacaggaagt 780
gattgggaag ttcaaataat tcttcaccgc aacgacccat ataacattga gtacttccca 840
tcgcgagtca gggatatgca gtctttgatt tataaactct gccgtgcgaa ctacacattc 900
aatttgctct ctcacgtatt caatgttcac aaggggatca aagaagatga tacaatgtac 960
tcgaaagtcg tcactgctca cacaaagcaa ttttggaaaa tgaggtattt atttttttgt 1020
tgtagagaat tcccaagata tgcttgtgaa tttacagaac gctttcccgt tacactgccg 1080
aaatcgacaa gcagtaccca gacactacaa caagataatt tgccagatgt ttccttattt 1140
ttttcaggag tattcagaat gttcacgcaa ttctcgaaat tttcagagca tttgaacatt 1200
tttaaagccg gaaaagcgta ctgttttgtt gtttctgtca ctttctggt gtctttaaaa 1260
tatggagaaa aataa 1275

<210> 82
<211> 424
<212> PRT
<213> Caenorhabditis elegans

<400> 82
Met Cys Thr Phe Lys Lys Phe Asp Gly Glu Thr Arg Lys Thr Arg Ile
1 5 10 15

Gln Ile Leu Tyr Phe Ala Ala Ser Leu Val Asn Leu Asp Leu Lys Pro
20 25 30

Val Lys Leu Asn Ser Asn Ala Asn Ile Cys Val Lys Ile Glu Thr Ser
35 40 45

His Phe Thr Ser Gly Thr Tyr Tyr Ile Asn Leu Ala Ser Val Gln Phe
50 55 60

Lys Gly Asn Ala Pro Gly Ser Asp Ala Glu Gly Arg Phe Phe Lys Lys
65 70 75 80

Leu His Gly Lys Pro Glu Asn Asn Tyr Asn Ser Leu Gln Thr Thr Val
85 90 95

His Phe Ala Phe Arg Ile Ser Pro Ser Gln Thr Glu Cys Pro Val Ile
100 105 110

Tyr Thr Ser Gly Tyr Lys Asp Cys Val Thr Phe Phe Gln Lys Asn Thr
115 120 125

Glu Leu Leu Glu Glu Met Glu Asp Pro Phe Gln Ile Tyr Pro Ile Asn
130 135 140

Leu Met Arg Asn Ile Ala Arg Arg Gly Ala Lys Ser Asp Leu His Leu
145 150 155 160

Ile Val Asp Thr Asp Met Val Met Ser Thr Asn Phe Ala Lys Met Val
165 170 175

Lys Pro Val Ala Asn Arg Met Ile Asp Gly Met Asn Lys Gln Val Leu
180 185 190

Val Val Arg Arg Phe Glu Thr Asn Glu Thr Glu Leu Pro Leu Asn Leu
195 200 205

Asp Glu Leu Glu Gln Gly Leu Leu Asn Glu Asn Thr Phe Glu Phe His
210 215 220

His Ser Phe Phe Phe Val Gly His Gln Ile Pro Asn Leu Ser Glu Trp
225 230 235 240

Phe Glu Asn Ser Tyr Ala Ser Glu Glu Thr Thr Ala Trp Glu Ile Pro
245 250 255

Tyr Thr Gly Ser Asp Trp Glu Val Gln Ile Ile Leu His Arg Asn Asp
260 265 270

Pro Tyr Asn Ile Glu Tyr Phe Pro Ser Arg Val Arg Asp Met Gln Ser
275 280 285

Leu Ile Tyr Lys Leu Cys Arg Ala Asn Tyr Thr Phe Asn Leu Leu Ser
290 295 300

His Val Phe Asn Val His Lys Gly Ile Lys Glu Asp Asp Thr Met Tyr
305 310 315 320

Ser Lys Val Val Thr Ala His Thr Lys Gln Phe Trp Lys Met Arg Tyr
325 330 335

Leu Phe Phe Cys Cys Arg Glu Phe Pro Arg Tyr Ala Cys Glu Phe Thr
340 345 350

Glu Arg Phe Pro Val Thr Leu Pro Lys Ser Thr Ser Ser Thr Gln Thr
355 360 365

Leu Gln Gln Asp Asn Leu Pro Asp Val Ser Leu Phe Phe Ser Gly Val
370 375 380

Phe Arg Met Phe Thr Gln Phe Ser Lys Phe Ser Glu His Leu Asn Ile
385 390 395 400

120
Phe Lys Ala Gly Lys Ala Tyr Cys Phe Val Val Ser Val Thr Phe Leu
405 410 415

Val Ser Leu Lys Tyr Gly Glu Lys
420

<210> 83
<211> 370
<212> PRT
<213> Caenorhabditis elegans

<400> 83
Met Glu Asp Asp Thr Pro Asp Val Ser Ser Asp Ser Asn Gly Asp Ala
1 5 10 15

Ala Tyr Ser Asp Tyr Phe Leu Asp Tyr Lys Ser Ile Met Asp Glu Ile
20 25 30

Thr Ile Thr Thr Gln Pro Lys Ser Gly Tyr Val Ile Arg Asn Lys Pro
35 40 45

Leu Arg Leu Gln Cys Arg Ala Asn His Ala Thr Lys Ile Arg Tyr Lys
50 55 60

Cys Ser Ser Lys Trp Ile Asp Asp Ser Arg Ile Glu Lys Leu Ile Gly
65 70 75 80

Thr Asp Ser Thr Ser Gly Val Gly Tyr Ile Asp Ala Ser Val Asp Ile
85 90 95

Ser Arg Ile Asp Val Asp Thr Ser Gly His Val Asp Ala Phe Gln Cys
100 105 110

Gln Cys Tyr Ala Ser Gly Asp Asp Asp Gln Asp Val Val Ala Ser Asp
115 120 125

Val Ala Thr Val His Leu Ala Tyr Met Arg Lys His Phe Leu Lys Ser
130 135 140

Pro Val Ala Gln Arg Val Gln Glu Gly Thr Thr Leu Gln Leu Pro Cys
145 150 155 160

Gln Ala Pro Glu Ser Asp Pro Lys Ala Glu Leu Thr Trp Tyr Lys Asp
165 170 175

Gly Val Val Val Gln Pro Asp Ala Asn Val Ile Arg Ala Ser Asp Gly
180 185 190

Ser Leu Ile Met Ser Ala Ala Arg Leu Ser Asp Ser Gly Asn Tyr Thr
195 200 205

Cys Glu Ala Thr Asn Val Ala Asn Ser Arg Lys Thr Asp Pro Val Glu
210 215 220

Val Gln Ile Tyr Val Asp Gly Gly Trp Ser Glu Trp Ser Pro Trp Ile
225 230 235 240

Gly Thr Cys His Val Asp Cys Pro Leu Leu Arg Gln His Ala His Arg
245 250 255

Ile Arg Asp Pro His Asp Val Leu Pro His Gln Arg Arg Thr Arg Thr
260 265 270

Cys Asn Asn Pro Ala Pro Leu Asn Asp Gly Glu Tyr Cys Lys Gly Glu
275 280 285

Glu Glu Met Thr Arg Ser Cys Lys Val Pro Cys Lys Leu Asp Gly Gly
290 295 300

Trp Ser Ser Trp Ser Asp Trp Ser Ala Cys Ser Ser Ser Cys His Arg
305 310 315 320

Tyr Arg Thr Arg Ala Cys Thr Val Pro Pro Pro Met Asn Gly Gly Gln
325 330 335

Pro Cys Phe Gly Asp Asp Leu Met Thr Gln Glu Cys Pro Ala Gln Leu
340 345 350

Cys Thr Ala Asp Ser Ser Arg Ile Val Ile Ser Asp Thr Ala Val Tyr
355 360 365

Gly Ser
370

<210> 84
<211> 20
<212> PRT
<213> Caenorhabditis elegans

<400> 84
Val Ala Ser Ile Phe Ile Val Ala Ser Phe Ile Leu Ala Ile Leu Ala
1 5 10 15

Met Phe Cys Cys
20

<210> 85
<211> 122
<212> PRT
<213> Caenorhabditis elegans

<400> 85
Lys Arg Gly Asn Ser Lys Lys Ser Lys Pro Leu Lys Pro Gln Lys Met
1 5 10 15

Asn Ser Glu Lys Ala Gly Gly Ile Tyr Tyr Ser Glu Pro Pro Gly Val
20 25 30

Arg Arg Leu Leu Leu Glu His Gln His Gly Thr Leu Leu Gly Glu Lys
35 40 45

Ile Ser Ser Cys Ser Gln Tyr Phe Glu Pro Pro Pro Leu Pro His Ser
50 55 60

Thr Thr Leu Arg Ser Gly Lys Ser Ala Phe Ser Gly Tyr Ser Ser Thr
65 70 75 80

Arg Asn Ala Gly Ser Arg Ala Ala Leu Ile Gln Glu Cys Ser Ser Ser
85 90 95

Ser Ser Gly Ser Gly Gly Lys Arg Thr Met Leu Arg Thr Ser Ser Ser
100 105 110

Asn Cys Ser Asp Asp Asp Asn Tyr Ala Thr
115 120

<210> 86
<211> 165
<212> PRT
<213> Caenorhabditis elegans

<400> 86
Leu Tyr Asp Tyr Met Glu Asp Lys Ser Val Leu Gly Leu Asp Thr Ser
1 5 10 15

Gln Asn Ile Val Ala Ala Gln Ile Asp Ser Asn Gly Ala Arg Leu Ser
20 25 30

Leu Ser Lys Ser Gly Ala Arg Leu Ile Val Pro Glu Leu Ala Val Glu
35 40 45

Gly Glu Lys Met Leu Tyr Leu Ala Val Ser Asp Thr Leu Thr Asp Gln
50 55 60

Pro His Leu Lys Pro Ile Glu Ser Ala Leu Ser Pro Val Ile Val Ile
65 70 75 80

Gly Gln Cys Asp Val Ser Met Ser Ala His Asp Asn Ile Leu Arg Arg
85 90 95

Pro Val Val Val Ser Phe Arg His Cys Ala Ser Thr Phe Pro Arg Asp
100 105 110

Asn Trp Gln Phe Thr Leu Tyr Ala Asp Glu Gly Ser Gly Trp Gln Lys
115 120 125

Ala Val Thr Ile Gly Glu Glu Asn Leu Asn Thr Asn Met Phe Val Gln
130 135 140

Phe Glu Gln Pro Gly Lys Lys Asn Asp Gly Phe Gly Trp Cys His Val
145 150 155 160

Met Thr Tyr Ser Leu
165

<210> 87
<211> 157
<212> PRT
<213> Caenorhabditis elegans

<400> 87
Ala Arg Leu Met Leu Ala Gly His Pro Arg Arg Asn Ser Leu Ser Ala
1               5                   10                  15

Ala Lys Arg Val His Leu Ala Val Phe Gly Pro Thr Glu Met Ser Ala
            20                  25                  30

Tyr Arg Arg Pro Phe Glu Leu Arg Val Tyr Cys Val Pro Glu Thr Gly
        35                  40                  45

Ala Ala Met Glu Ser Val Trp Lys Gln Glu Asp Gly Ser Arg Leu Leu
    50                  55                  60

Cys Glu Ser Asn Asp Phe Ile Leu Asn Glu Lys Gly Asn Leu Cys Ile
65                  70                  75                  80

Cys Ile Glu Asp Val Ile Pro Gly Phe Ser Cys Asp Gly Pro Glu Val
                85                  90                  95

Val Glu Ile Ser Glu Thr Gln His Arg Phe Val Ala Gln Asn Gly Leu
            100                 105                 110

His Cys Ser Leu Lys Phe Arg Pro Lys Glu Ile Asn Gly Ser Gln Phe
        115                 120                 125

Ser Thr Arg Val Ile Val Tyr Gln Lys Ala Ser Ser Thr Glu Pro Met
    130                 135                 140

Val Met Glu Val Ser Asn Glu Pro Glu Leu Tyr Asp Ala
145                 150                 155

<210> 88
<211> 113
<212> PRT
<213> Caenorhabditis elegans

<400> 88
Thr Ser Glu Glu Arg Glu Lys Gly Ser Val Cys Val Glu Phe Arg Leu
1               5                   10                  15

Pro Phe Gly Val Lys Asp Glu Leu Ala Arg Leu Leu Asp Met Pro Asn
            20                  25                  30

Glu Ser His Ser Asp Trp Arg Gly Leu Ala Lys Lys Leu His Tyr Asp
        35                  40                  45

Arg Tyr Leu Gln Phe Phe Ala Ser Phe Pro Asp Cys Ser Pro Thr Ser
    50                  55                  60

Leu Leu Leu Asp Leu Trp Glu Ala Ser Ser Ser Gly Ser Ala Arg Ala
65                  70                  75                  80

Val Pro Asp Leu Leu Gln Thr Leu Arg Val Met Gly Arg Pro Asp Ala
                85                  90                  95

Val Met Val Leu Glu Arg Phe Leu Ser Ala Phe Pro Gln Ile Val Ser

100 105 110

Pro

<210> 89
<211> 437
<212> PRT
<213> Homo sapiens

<400> 89
His Met Ala Thr Leu His His Ser Ser Pro Thr Ser Glu Ala Glu Glu
1 5 10 15

Phe Val Ser Arg Leu Ser Thr Gln Asn Tyr Phe Arg Ser Leu Pro Arg
20 25 30

Gly Thr Ser Asn Met Thr Tyr Gly Thr Phe Asn Phe Leu Gly Gly Arg
35 40 45

Leu Met Ile Pro Asn Thr Gly Ile Ser Leu Leu Ile Pro Pro Asp Ala
50 55 60

Ile Pro Arg Gly Lys Ile Tyr Glu Ile Tyr Leu Thr Leu His Lys Pro
65 70 75 80

Glu Asp Val Arg Leu Pro Leu Ala Gly Cys Gln Thr Leu Leu Ser Pro
85 90 95

Ile Val Ser Cys Gly Pro Pro Gly Val Leu Leu Thr Arg Pro Val Ile
100 105 110

Leu Ala Met Asp His Cys Gly Glu Pro Ser Pro Asp Ser Trp Ser Leu
115 120 125

Arg Leu Lys Lys Gln Ser Cys Glu Gly Ser Trp Glu Asp Val Leu His
130 135 140

Leu Gly Glu Glu Ala Pro Ser His Leu Tyr Tyr Cys Gln Leu Glu Ala
145 150 155 160

Ser Ala Cys Tyr Val Phe Thr Glu Gln Leu Gly Arg Phe Ala Leu Val
165 170 175

Gly Glu Ala Leu Ser Val Ala Ala Ala Lys Arg Leu Lys Leu Leu Leu
180 185 190

Phe Ala Pro Val Ala Cys Thr Ser Leu Glu Tyr Asn Ile Arg Val Tyr
195 200 205

Cys Leu His Asp Thr His Asp Ala Leu Lys Glu Val Val Gln Leu Glu
210 215 220

Lys Gln Leu Gly Gly Gln Leu Ile Gln Glu Pro Arg Val Leu His Phe
225 230 235 240

Lys Asp Ser Tyr His Asn Leu Arg Leu Ser Ile His Asp Val Pro Ser

245 250 255

Ser Leu Trp Lys Ser Lys Leu Leu Val Ser Tyr Gln Glu Ile Pro Phe
260 265 270

Tyr His Ile Trp Asn Gly Thr Gln Arg Tyr Leu His Cys Thr Phe Thr
275 280 285

Leu Glu Arg Val Ser Pro Ser Thr Ser Asp Leu Ala Cys Lys Leu Trp
290 295 300

Val Trp Gln Val Glu Gly Asp Gly Gln Ser Phe Ser Ile Asn Phe Asn
305 310 315 320

Ile Thr Lys Asp Thr Arg Phe Ala Glu Leu Leu Ala Leu Glu Ser Glu
325 330 335

Ala Gly Val Gln Ala Leu Val Gly Pro Ser Ala Phe Lys Ile Pro Phe
340 345 350

Leu Ile Arg Gln Lys Ile Ile Ser Ser Leu Asp Pro Pro Cys Arg Arg
355 360 365

Gly Ala Asp Trp Arg Thr Leu Ala Gln Lys Leu His Leu Asp Ser His
370 375 380

Leu Ser Phe Phe Ala Ser Lys Pro Ser Pro Thr Ala Met Ile Leu Asn
385 390 395 400

Leu Trp Glu Ala Arg His Phe Pro Asn Gly Asn Leu Ser Gln Leu Ala
405 410 415

Ala Ala Val Ala Gly Leu Gly Gln Pro Asp Ala Gly Leu Phe Thr Val
420 425 430

Ser Glu Ala Glu Cys
435

<210> 90
<211> 931
<212> PRT
<213> Homo sapiens

<400> 90
Met Arg Lys Gly Leu Arg Ala Thr Ala Ala Arg Cys Gly Leu Gly Leu
1 5 10 15

Gly Tyr Leu Leu Gln Met Leu Val Leu Pro Ala Leu Ala Leu Leu Ser
20 25 30

Ala Ser Gly Thr Gly Ser Ala Ala Gln Asp Asp Asp Phe Phe His Glu
35 40 45

Leu Pro Glu Thr Phe Pro Ser Asp Pro Pro Glu Pro Leu Pro His Phe
50 55 60

Leu Ile Glu Pro Glu Glu Ala Tyr Ile Val Lys Asn Lys Pro Val Asn

65 70 75 80

Leu Tyr Cys Lys Ala Ser Pro Ala Thr Gln Ile Tyr Phe Lys Cys Asn
85 90 95

Ser Glu Trp Val His Gln Lys Asp His Ile Val Asp Glu Arg Val Asp
100 105 110

Glu Thr Ser Gly Leu Ile Val Arg Glu Val Ser Ile Glu Ile Ser Arg
115 120 125

Gln Gln Val Glu Glu Leu Phe Gly Pro Glu Asp Tyr Trp Cys Gln Cys
130 135 140

Val Ala Trp Ser Ser Ala Gly Thr Thr Lys Ser Arg Lys Ala Tyr Val
145 150 155 160

Arg Ile Ala Tyr Leu Arg Lys Thr Phe Glu Gln Glu Pro Leu Gly Lys
165 170 175

Glu Val Ser Leu Glu Gln Glu Val Leu Leu Gln Cys Arg Pro Pro Glu
180 185 190

Gly Ile Pro Val Ala Glu Val Glu Trp Leu Lys Asn Glu Asp Ile Ile
195 200 205

Asp Pro Val Glu Asp Arg Asn Phe Tyr Ile Thr Ile Asp His Asn Leu
210 215 220

Ile Ile Lys Gln Ala Arg Leu Ser Asp Thr Ala Asn Tyr Thr Cys Val
225 230 235 240

Ala Lys Asn Ile Val Ala Lys Arg Lys Ser Thr Thr Ala Thr Val Ile
245 250 255

Val Tyr Val Asn Gly Gly Trp Ser Thr Trp Thr Glu Trp Ser Val Cys
260 265 270

Asn Ser Arg Cys Gly Arg Gly Tyr Gln Lys Arg Thr Arg Thr Cys Thr
275 280 285

Asn Pro Ala Pro Leu Asn Gly Gly Ala Phe Cys Glu Gly Gln Ser Val
290 295 300

Gln Lys Ile Ala Cys Thr Thr Leu Cys Pro Val Asp Gly Arg Trp Thr
305 310 315 320

Pro Trp Ser Lys Trp Ser Thr Cys Gly Thr Glu Cys Thr His Trp Arg
325 330 335

Arg Arg Glu Cys Thr Ala Pro Ala Pro Lys Asn Gly Gly Lys Asp Cys
340 345 350

Asp Gly Leu Val Leu Gln Ser Lys Asn Cys Thr Asp Gly Leu Cys Met
355 360 365

Gln Thr Ala Pro Asp Ser Asp Asp Val Ala Leu Tyr Val Gly Ile Val
370 375 380

Ile Ala Val Ile Val Cys Leu Ala Ile Ser Val Val Val Ala Leu Phe
385 390 395 400

Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile Asp Ser
405 410 415

Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala Ala Arg
420 425 430

Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala Ala Met
435 440 445

Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys Ile Pro
450 455 460

Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys Ile Lys
465 470 475 480

Val Tyr Asn Thr Ser Gly Ala Val Ser Pro Gln Asp Asp Leu Ser Glu
485 490 495

Phe Thr Ser Lys Leu Ser Pro Gln Met Thr Gln Ser Leu Leu Glu Asn
500 505 510

Glu Ala Leu Ser Leu Lys Asn Gln Ser Leu Ala Arg Gln Thr Asp Pro
515 520 525

Ser Cys Thr Ala Phe Gly Ser Phe Asn Ser Leu Gly Gly His Leu Ile
530 535 540

Val Pro Asn Ser Gly Val Ser Leu Leu Ile Pro Ala Gly Ala Ile Pro
545 550 555 560

Gln Gly Arg Val Tyr Glu Met Tyr Val Thr Val His Arg Lys Glu Thr
565 570 575

Met Arg Pro Pro Met Asp Asp Ser Gln Thr Leu Leu Thr Pro Val Val
580 585 590

Ser Cys Gly Pro Pro Gly Ala Leu Leu Thr Arg Pro Val Val Leu Thr
595 600 605

Met His His Cys Ala Asp Pro Asn Thr Glu Asp Trp Lys Ile Leu Leu
610 615 620

Lys Asn Gln Ala Ala Gln Gly Gln Trp Glu Asp Val Val Val Val Gly
625 630 635 640

Glu Glu Asn Phe Thr Thr Pro Cys Tyr Ile Lys Leu Asp Ala Glu Ala
645 650 655

Cys His Ile Leu Thr Glu Asn Leu Ser Thr Tyr Ala Leu Val Gly His
660 665 670

Ser Thr Thr Lys Ala Ala Ala Lys Arg Leu Lys Leu Ala Ile Phe Gly
675 680 685

Pro Leu Cys Cys Ser Ser Leu Glu Tyr Ser Ile Arg Val Tyr Cys Leu
690 695 700

Asp Asp Thr Gln Asp Ala Leu Lys Glu Ile Leu His Leu Glu Arg Gln
705 710 715 720

Thr Gly Gly Gln Leu Leu Glu Glu Pro Lys Ala Leu His Phe Lys Gly
725 730 735

Ser Thr His Asn Leu Arg Leu Ser Ile His Asp Ile Ala His Ser Leu
740 745 750

Trp Lys Ser Lys Leu Leu Ala Lys Tyr Gln Glu Ile Pro Phe Tyr His
755 760 765

Val Trp Ser Gly Ser Gln Arg Asn Leu His Cys Thr Phe Thr Leu Glu
770 775 780

Arg Phe Ser Leu Asn Thr Val Glu Leu Val Cys Lys Leu Cys Val Arg
785 790 795 800

Gln Val Glu Gly Glu Gly Gln Ile Phe Gln Leu Asn Cys Thr Val Ser
805 810 815

Glu Glu Pro Thr Gly Ile Asp Leu Pro Leu Leu Asp Pro Ala Asn Thr
820 825 830

Ile Thr Thr Val Thr Gly Pro Ser Ala Phe Ser Ile Pro Leu Pro Ile
835 840 845

Arg Gln Lys Leu Cys Ser Ser Leu Asp Ala Pro Gln Thr Arg Gly His
850 855 860

Asp Trp Arg Met Leu Ala His Lys Leu Asn Leu Asp Arg Tyr Leu Asn
865 870 875 880

Tyr Phe Ala Thr Lys Ser Ser Pro Thr Gly Val Ile Leu Asp Leu Trp
885 890 895

Glu Ala Gln Asn Phe Pro Asp Gly Asn Leu Ser Met Leu Ala Ala Val
900 905 910

Leu Glu Glu Met Gly Arg His Glu Thr Val Val Ser Leu Ala Ala Glu
915 920 925

Gly Gln Tyr
930

<210> 91
<211> 9700
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: plasmid
pGC1037

<400> 91
tatggccaca ctgcaccaca gctctcccac ctctgaggcc gaggagttcg tctcccgcct 60
ctccacccag aactacttcc gctccctgcc ccgaggcacc agcaacatga cctatgggac 120
cttcaacttc ctcgggggcc ggctgatgat ccctaataca ggaatcagcc tcctcatccc 180
cccagatgcc ataccccgag ggaagatcta tgagatctac ctcacgctgc acaagccgga 240
agacgtgagg ttgcccctag ctggctgtca gaccctgctg agtcccatcg ttagctgtgg 300
accccctggc gtcctgctca cccggccagt catcctggct atggaccact gtggggagcc 360
cagccctgac agctggagcc tgcgcctcaa aaagcagtcg tgcgagggca gctgggagga 420
tgtgctgcac ctgggcgagg aggcgccctc ccacctctac tactgccagc tggaggccag 480
tgcctgctac gtcttcaccg agcagctggg ccgctttgcc ctggtgggag aggccctcag 540
cgtggctgcc gccaagcgcc tcaagctgct tctgtttgcg ccggtggcct gcacctccct 600
cgagtacaac atccgggtct actgcctgca tgacacccac gatgcactca aggaggtggt 660
gcagctggag aagcagctgg ggggacagct gatccaggag ccacgggtcc tgcacttcaa 720
ggacagttac cacaacctgc gcctatccat ccacgatgtg cccagctccc tgtggaagag 780
taagctcctt gtcagctacc aggagatccc cttttatcac atctggaatg gcacgcagcg 840
gtacttgcac tgcaccttca ccctggagcg tgtcagcccc agcactagtg acctggcctg 900
caagctgtgg gtgtggcagg tggagggcga cgggcagagc ttcagcatca acttcaacat 960
caccaaggac acaaggtttg ctgagctgct ggctctggag agtgaagcgg gggtccaagc 1020
cctggtgggc cccagtgcct tcaagatccc cttcctcatt cggcagaaga taatttccag 1080
cctggaccca ccctgtaggc ggggtgccga ctggcggact ctggcccaga aactccacct 1140
ggacagccat ctcagcttct ttgcctccaa gcccagcccc acagccatga tcctcaacct 1200
gtgggaggcg cggcacttcc ccaacggcaa cctcagccag ctggctgcag cagtggctgg 1260
actgggccag ccagacgctg gcctcttcac agtgtcggag gctgagtgct gagtcgacct 1320
gcagccaagc taattccggg cgaatttctt atgatttatg atttttatta ttaaataagt 1380
tataaaaaaa ataagtgtat acaaatttta aagtgactct taggttttaa aacgaaaatt 1440
cttgttcttg agtaactctt tcctgtaggt caggttgctt tctcaggtat agcatgaggt 1500
cgctcttatt gaccacacct ctaccggcat gcaagcttgg cgtaatcatg gtcatagctg 1560
tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata 1620
aagtgtaaag cctggggtgc ctaatgagtg aggtaactca cattaattgc gttgcgctca 1680
ctgcccgctt tccagtcggg aaacctgtcg tgccagctgg attaatgaat cggccaacgc 1740
gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 1800
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 1860
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 1920
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 1980
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 2040
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 2100
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 2160
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 2220
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 2280
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 2340
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 2400
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 2460
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 2520
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 2580
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 2640
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 2700
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 2760
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 2820
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 2880
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 2940
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 3000
agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt 3060
atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg 3120
tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 3180
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 3240
agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg 3300
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 3360
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 3420

ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt 3480
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 3540
ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 3600
atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 3660
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgaacgaag catctgtgct 3720
tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag 3780
ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac gaagaatctg 3840
tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat 3900
ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta ccaacaaaga 3960
atctatactt cttttttgtt ctacaaaaat gcatcccgag agcgctattt ttctaacaaa 4020
gcatcttaga ttactttttt tctcctttgt gcgctctata atgcagtctc ttgataactt 4080
tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat tttctcttcc 4140
ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc tgcgggtgca 4200
ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt gcgcatactt 4260
tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta tgaacggttt 4320
cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt attgttttcg 4380
attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat actagagata 4440
aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa ggtggatggg 4500
taggttatat agggatatag cacagagata tatagcaaag agatactttt gagcaatgtt 4560
tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg tttttggttt 4620
tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag ttcctatact 4680
ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa cgagcgcttc 4740
cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac ctatatctgc 4800
gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt atgcttaaat 4860
gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct cctgtgatat 4920
tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta gctgttctat 4980
atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt cctttgatat 5040
tggatcatat taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 5100
gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 5160
cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 5220
cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat 5280
tgtactgaga gtgcaccata gatcaacgac attactatat atataatata ggaagcattt 5340
aatagacagc atcgtaatat atgtgtactt tgcagttatg acgccagatg gcagtagtgg 5400
aagatattct ttattgaaaa atagcttgtc accttacgta caatcttgat ccggagcttt 5460
tctttttttg ccgattaaga attaattcgg tcgaaaaaag aaaaggagag ggccaagagg 5520
gagggcattg gtgactattg agcacgtgag tatacgtgat taagcacaca aaggcagctt 5580
ggagtatgtc tgttattaat ttcacaggta gttctggtcc attggtgaaa gtttgcggct 5640
tgcagagcac agaggccgca gaatgtgctc tagattccga tgctgacttg ctgggtatta 5700
tatgtgtgcc caatagaaag agaacaattg acccggttat tgcaaggaaa atttcaagtc 5760
ttgtaaaagc atataaaaat agttcaggca ctccgaaata cttggttggc gtgtttcgta 5820
atcaacctaa ggaggatgtt ttggctctgg tcaatgatta cggcattgat atcgtccaac 5880
tgcatggaga tgagtcgtgg caagaatacc aagagttcct cggtttgcca gttattaaaa 5940
gactcgtatt tccaaaagac tgcaacatac tactcagtgc agcttcacag aaacctcatt 6000
cgtttattcc cttgtttgat tcagaagcag gtgggacagg tgaacttttg gattggaact 6060
cgatttctga ctgggttgga aggcaagaga gccccgaaag cttacatttt atgttagctg 6120
gtggactgac gccagaaaat gttggtgatg cgcttagatt aaatggcgtt attggtgttg 6180
atgtaagcgg aggtgtggag acaaatggtg taaaagactc taacaaaata gcaaatttcg 6240
tcaaaaatgc taagaaatag gttattactg agtagtattt atttaagtat tgtttgtgca 6300
cttgccgatc tatgcggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca 6360
ggaaattgta aacgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc 6420
atttttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 6480
gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc 6540
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc 6600
ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaagggag 6660
cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa 6720
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 6780
cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcg cgccattcgc cattcaggct 6840
gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa 6900

agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 6960
ttgtaaaacg acggccagtc gtccaagctt tcgcgagctc gagatcccga gctttgcaaa 7020
ttaaagcctt cgagcgtccc aaaaccttct caagcaaggt tttcagtata atgttacatg 7080
cgtacacgcg tctgtacaga aaaaaaagaa aaatttgaaa tataaataac gttcttaata 7140
ctaacataac tataaaaaaa taaataggga cctagacttc aggttgtcta actccttcct 7200
tttcggttag agcggatgtg gggggagggc gtgaatgtaa gcgtgacata actaattaca 7260
tgatatcgac aaaggaaaag gggcctgttt actcacaggc tttttcaag taggtaatta 7320
agtcgtttct gtcttttcc ttcttcaacc caccaaaggc catcttggta cttttttttt 7380
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 7440
tttttttttt tttttttttt tttttcatag aaataataca gaagtagatg ttgaattaga 7500
ttaaactgaa gatatataat ttattggaaa atacatagag cttttgttg atgcgcttaa 7560
gcgatcaatt caacaacacc accagcagct ctgatttttt cttcagccaa cttggagacg 7620
aatctagctt tgacgataac tggaacattt ggaattctac ccttacccaa gatcttaccg 7680
taaccggctg ccaaagtgtc aataactgga gcagtttcct tagaagcaga tttcaagtat 7740
tggtctctct tgtcttctgg gatcaatgtc cacaatttgt ccaagttcaa gactggcttc 7800
tagaaatgag cttgttgctt gtggaagtat ctcataccaa nccttaccga aataacctgg 7860
atggtattta tccatgttaa ttctgtggtg atgttgacca ccggccatac ctctaccacc 7920
gggtgcttt ctgtgcttac cgatacgacc tttaccggct gagacgtgac ctctgtgctt 7980
tctagtctta gtgaatctgg aaggcattct tgattagttg gatgattgtt ctgggattta 8040
atgcaaaaaa atcactaaga aggaaaaaaa tcaacggaga aagcaaacgc catcttaaat 8100
atacgggata cagatgaaag gtttgaacct atctgggaaa atacgcatta aacaagcgaa 8160
aaactgcgag gaaaattgtt tgcgtctctg cgggctattc acgcgccaga ggaaaatagg 8220
aaaaataaca gggcattaga aaaataattt tgattttggt aatgtgtggg tccctggtgt 8280
acagatgtta cattggttac agtactcttg tttttgctgt gtttttcgat gaatctccaa 8340
aatggttgtt agcacatgga agagtcaccg atgctaagtt atctctatgt aagctacgtg 8400
gcgtgacttt tgatgaagcc gcacaagaga tacaggattg gcaactgcaa atagaatctg 8460
gggatctaga tatccttttg ttgtttccgg gtgtacaata tggacttcct cttttctggc 8520
aaccaaaccc atacatcggg attcctataa taccttcgtt ggtctcccta acatgtaggt 8580
ggcggagggg agatatacaa tagaacagat accagacaag acataatggg ctaaacaaga 8640
ctacaccaat tacactgcct cattgatggt ggtacataac gaactaatac tgtagcccta 8700
gacttgatag ccatcatcat atcgaagttt cactaccctt tttccatttg ccatctattg 8760
aagtaataat aggcgcatgc aacttctttt cttttttttt cttttctctc tcccccgttg 8820
ttgtctcacc atatccgcaa tgacaaaaaa aatgatggaa gacactaaag gaaaaaatta 8880
acgacaaaga cagcaccaac agatgtcgtt gttccagagc tgatgagggg tatcttcgaa 8940
cacacgaaac tttttccttc cttcattcac gcacactact ctctaatgag caacggtata 9000
cggccttcct tccagttact tgaatttgaa ataaaaaaag tttgccgctt tgctatcaag 9060
tataaataga cctgcaatta ttaatctttt gtttcctcgt cattgttctc gttccctttc 9120
ttccttgttt ctttttctgc acaatatttc aagctatacc aagcatacaa tcaactccaa 9180
gcttgaagca agcctcctga aagatgaagc tactgtcttc tatcgaacaa gcatgcgata 9240
tttgccgact taaaaagctc aagtgctcca aagaaaaacc gaagtgcgcc aagtgtctga 9300
agaacaactg ggagtgtcgc tactctccca aaaccaaaag gtctccgctg actagggcac 9360
atctgacaga agtggaatca aggctagaaa gactggaaca gctatttcta ctgatttttc 9420
ctcgagaaga ccttgacatg attttgaaaa tggattcttt acaggatata aaagcattgt 9480
taacaggatt atttgtacaa gataatgtga ataaagatgc cgtcacagat agattggctt 9540
cagtggagac tgatatgcct ctaacattga gacagcatag aataagtgcg acatcatcat 9600
cggaagagag tagtaacaaa ggtcaaagac agttgactgt atcgccggaa ttcatggctt 9660
acccatacga tgttccagat tacgctagct tgggtggtca 9700

<210> 92
<211> 9749
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: plasmid
pGC1003

<400> 92

```
agaatctggg gatctagata tccttttgtt gtttccgggt gtacaatatg gacttcctct 60
tttctggcaa ccaaacccat acatcgggat tcctataata ccttcgttgg tctccctaac 120
atgtaggtgg cggaggggag atatacaata gaacagatac cagacaagac ataatgggct 180
aaacaagact acaccaatta cactgcctca ttgatggtgg tacataacga actaatactg 240
tagccctaga cttgatagcc atcatcatat cgaagtttca ctaccctttt tccatttgcc 300
atctattgaa gtaataatag gcgcatgcaa cttcttttct tttttttct tttctctctc 360
ccccgttgtt gtctcaccat atccgcaatg acaaaaaaaa tgatggaaga cactaaagga 420
aaaaattaac gacaaagaca gcaccaacag atgtcgttgt tccagagctg atgaggggta 480
tcttcgaaca cacgaaactt tttccttcct tcattcacgc acactactct ctaatgagca 540
acggtatacg gccttccttc cagttacttg aatttgaaat aaaaaaagtt tgccgctttg 600
ctatcaagta taaatagacc tgcaattatt aatcttttgt ttcctcgtca ttgttctcgt 660
tccctttctt ccttgtttct ttttctgcac aatatttcaa gctataccaa gcatacaatc 720
aactccaagc ttgaagcaag cctcctgaaa gatgaagcta ctgtcttcta tcgaacaagc 780
atgcgatatt tgccgactta aaaagctcaa gtgctccaaa gaaaaaccga agtgcgccaa 840
gtgtctgaag aacaactggg agtgtcgcta ctctcccaaa accaaaaggt ctccgctgac 900
tagggcacat ctgacagaag tggaatcaag gctagaaaga ctggaacagc tatttctact 960
gatttttcct cgagaagacc ttgacatgat tttgaaaatg gattctttac aggatataaa 1020
agcattgtta acaggattat ttgtacaaga taatgtgaat aaagatgccg tcacagatag 1080
attggcttca gtggagactg atatgcctct aacattgaga cagcatagaa taagtgcgac 1140
atcatcatcg gaagagagta gtaacaaagg tcaaagacag ttgactgtat cgccggaatt 1200
catggcttac ccatacgatg ttccagatta cgctagcttg ggtggtcata tggccatgga 1260
gtgctgtaaa cgtggcaatt caaaaaagtc gaaaccattg aaaccgcaga aaatgaattc 1320
ggaaaaagca ggaggaatct actattctga gcctccaggt gttcgaagat tacttttgga 1380
gcatcaacat ggaacattgc tcggagagaa aatatcatcg tgttctcaat attttgagcc 1440
tcctcctcta ccacattcaa caactcttcg atctggaaaa agtgcattca gtggatattc 1500
ttcaacaaga aacgctggct cacgagctgc tttaattcaa gaatgtagtt cgagtagttc 1560
aggatctggt ggaaaacgga caatgcttag aacttcaagc tcaaattgct cggatgacga 1620
taattatgct acgttatatg attatatgga agacaagagt gttcttggat tagacacttc 1680
acaaaatatt gttgctgcgc agattgatag caatggagca cgattgtctt tatcaaaaag 1740
tggagcaaga ttaattgtac ctgaattggc agtggaagga gagaaaatgc tttatctggc 1800
tgtttcggac accttaacag atcaacctca cctcaaacca atcgaatccg ctctctcccc 1860
agtcatcgtc atcggacaat gtgatgtttc aatgtcagct catgataata ttctccgtcg 1920
tcctgttgtt gtatccttcc gtcattgtgc ttcaactttt ccacgtgaca attggcaatt 1980
cacattgtat gctgatgaag gatccggatg gcagaaggct gtaacaattg gagaagagaa 2040
cttgaatacc aacatgtttg tacaattcga gcaacctggg aagaagaatg atggatttgg 2100
atggtgccac gttatgacat attccttagc gaggttaatg ctagctggtc atccgaggag 2160
aaattctttg agtgctgcaa aaagagttca tttagctgta tttggtccaa ctgaaatgtc 2220
agcttatcga agaccatttg aattgagagt ttattgtgtt cctgaaactg gagcagcaat 2280
ggaaagtgta tggaaacagg aagatggatc tcgtctactt tgtgaatcaa atgattttat 2340
tttaaatgaa aaaggaaatc tttgcatttg tattgaagat gttattcctg gattctcatg 2400
cgatggccct gaagttgtgg aaatctcgga aacacaacac agatttgttg ctcaaaatgg 2460
ccttcattgc tcattgaaat tccgaccaaa ggaaatcaac ggctcacagt tttcaacgag 2520
agtaattgtc tatcaaaagg caagctcaac ggagccaatg gtgatggaag tgagcaatga 2580
gccggagctc tatgatgcag ccccgggcg tcgacctgca gccaagctaa ttccgggcga 2640
atttcttatg atttatgatt tttattatta aataagttat aaaaaaaata agtgtataca 2700
aattttaaag tgactcttag gttttaaaac gaaaattctt gttcttgagt aactctttcc 2760
tgtaggtcag gttgctttct caggtatagc atgaggtcgc tcttattgac cacacctcta 2820
ccggcatgca agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc 2880
gctcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta 2940
atgagtgagg taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa 3000
cctgtcgtgc cagctggatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat 3060
tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg 3120
agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc 3180
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt 3240
gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag 3300
tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc 3360
cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc 3420
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt 3480
```

cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt 3540
atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc 3600
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa 3660
gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa 3720
gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg 3780
tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga 3840
agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg 3900
gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg 3960
aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt 4020
aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact 4080
ccccgtcgtg tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat 4140
gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg 4200
aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg 4260
ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat 4320
tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc 4380
ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt 4440
cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc 4500
agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga 4560
gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc 4620
gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa 4680
acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta 4740
acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg 4800
agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg 4860
aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat 4920
gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt 4980
tccccgaaaa gtgccacctg aacgaagcat ctgtgcttca ttttgtagaa caaaaatgca 5040
acgcgagagc gctaattttt caaacaaaga atctgagctg catttttaca gaacagaaat 5100
gcaacgcgaa agcgctattt taccaacgaa gaatctgtgc ttcatttttg taaaacaaaa 5160
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 5220
gaaatgcaac gcgagagcgc tattttacca acaaagaatc tatacttctt ttttgttcta 5280
caaaaatgca tcccgagagc gctatttttc taacaaagca tcttagatta cttttttttct 5340
cctttgtgcg ctctataatg cagtctcttg ataacttttt gcactgtagg tccgttaagg 5400
ttagaagaag gctactttgg tgtctatttt ctcttccata aaaaaagcct gactccactt 5460
cccgcgttta ctgattacta gcgaagctgc gggtgcattt tttcaagata aaggcatccc 5520
cgattatatt ctataccgat gtggattgcg catactttgt gaacagaaag tgatagcgtt 5580
gatgattctt cattggtcag aaaattatga acggtttctt ctattttgtc tctatatact 5640
acgtatagga aatgtttaca ttttcgtatt gttttcgatt cactctatga atagttctta 5700
ctacaatttt tttgtctaaa gagtaatact agagataaac ataaaaaatg tagaggtcga 5760
gtttagatgc aagttcaagg agcgaaaggt ggatgggtag gttatatagg gatatagcac 5820
agagatatat agcaaagaga tacttttgag caatgtttgt ggaagcggta ttcgcaatat 5880
tttagtagct cgttacagtc cggtgcgttt ttggtttttt gaaagtgcgt cttcagagcg 5940
cttttggttt tcaaaagcgc tctgaagttc ctatactttc tagagaatag gaacttcgga 6000
ataggaactt caaagcgttt ccgaaaacga gcgcttccga aaatgcaacg cgagctgcgc 6060
acatacagct cactgttcac gtcgcaccta tatctgcgtg ttgcctgtat atatatatac 6120
atgagaagaa cggcatagtg cgtgtttatg cttaaatgcg tacttatatg cgtctattta 6180
tgtaggatga aaggtagtct agtacctcct gtgatattat cccattccat gcggggtatc 6240
gtatgcttcc ttcagcacta ccctttagct gttctatatg ctgccactcc tcaattggat 6300
tagtctcatc cttcaatgct atcatttcct ttgatattgg atcatattaa gaaaccatta 6360
ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt ctcgcgcgtt 6420
tcggtgatga cggtgaaaac ctctgacaca tgcagctccc ggagacggtc acagcttgtc 6480
tgtaagcgga tgccgggagc agacaagccc gtcagggcgc gtcagcgggt gttggcgggt 6540
gtcggggctg gcttaactat gcggcatcag agcagattgt actgagagtg caccatagat 6600
caacgacatt actatatata taatatagga agcatttaat agacagcatc gtaatatatg 6660
tgtactttgc agttatgacg ccagatggca gtagtggaag atattcttta ttgaaaaata 6720
gcttgtcacc ttacgtacaa tcttgatccg gagcttttct ttttttgccg attaagaatt 6780
aattcggtcg aaaaaagaaa aggagagggc caagagggag ggcattggtg actattgagc 6840
acgtgagtat acgtgattaa gcacacaaag gcagcttgga gtatgtctgt tattaatttc 6900
acaggtagtt ctggtccatt ggtgaaagtt tgcggcttgc agagcacaga ggccgcagaa 6960

```
tgtgctctag attccgatgc tgacttgctg ggtattatat gtgtgcccaa tagaaagaga 7020
acaattgacc cggttattgc aaggaaaatt tcaagtcttg taaaagcata taaaaatagt 7080
tcaggcactc cgaaatactt ggttggcgtg tttcgtaatc aacctaagga ggatgttttg 7140
gctctggtca atgattacgg cattgatatc gtccaactgc atggagatga gtcgtggcaa 7200
gaataccaag agttcctcgg tttgccagtt attaaaagac tcgtatttcc aaaagactgc 7260
aacatactac tcagtgcagc ttcacagaaa cctcattcgt ttattccctt gtttgattca 7320
gaagcaggtg ggacaggtga acttttggat tggaactcga tttctgactg ggttggaagg 7380
caagagagcc ccgaaagctt acattttatg ttagctggtg gactgacgcc agaaaatgtt 7440
ggtgatgcgc ttagattaaa tggcgttatt ggtgttgatg taagcggagg tgtggagaca 7500
aatggtgtaa aagactctaa caaaatagca aatttcgtca aaaatgctaa gaaataggtt 7560
attactgagt agtatttatt taagtattgt ttgtgcactt gccgatctat gcggtgtgaa 7620
ataccgcaca gatgcgtaag gagaaaatac cgcatcagga aattgtaaac gttaatattt 7680
tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt ttttaaccaa taggccgaaa 7740
tcggcaaaat cccttataaa tcaaaagaat agaccgagat agggttgagt gttgttccag 7800
tttggaacaa gagtccacta ttaaagaacg tggactccaa cgtcaaaggg cgaaaaaccg 7860
tctatcaggg cgatggccca ctacgtgaac catcacccta atcaagtttt ttggggtcga 7920
ggtgccgtaa agcactaaat cggaacccta aagggagccc ccgatttaga gcttgacggg 7980
gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc gaaaggagcg ggcgctaggg 8040
cgctggcaag tgtagcggtc acgctgcgcg taaccaccac acccgccgcg cttaatgcgc 8100
cgctacaggg cgcgtcgcgc cattcgccat tcaggctgcg caactgttgg gaagggcgat 8160
cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat 8220
taagttgggt aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtcgtc 8280
caagctttcg cgagctcgag atcccgagct ttgcaaatta aagccttcga gcgtcccaaa 8340
accttctcaa gcaaggtttt cagtataatg ttacatgcgt acacgcgtct gtacagaaaa 8400
aaaagaaaaa tttgaaatat aaataacgtt cttaatacta acataactat aaaaaaataa 8460
atagggacct agacttcagg ttgtctaact ccttcctttt cggttagagc ggatgtgggg 8520
ggagggcgtg aatgtaagcg tgacataact aattacatga tatcgacaaa ggaaaagggg 8580
cctgtttact cacaggcttt tttcaagtag gtaattaagt cgtttctgtc tttttccttc 8640
ttcaacccac caaaggccat cttggtactt tttttttttt tttttttttt tttttttttt 8700
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 8760
ttcatagaaa taatacagaa gtagatgttg aattagatta aactgaagat atataattta 8820
ttggaaaata catagagctt tttgttgatg cgcttaagcg atcaattcaa caacaccacc 8880
agcagctctg atttttttctt cagccaactt ggagacgaat ctagctttga cgataactgg 8940
aacatttgga attctacccct tacccaagat cttaccgtaa ccggctgcca aagtgtcaat 9000
aactggagca gtttccttag aagcagattt caagtattgg tctctcttgt cttctgggat 9060
caatgtccac aatttgtcca agttcaagac tggcttccag aaatgagctt gttgcttgtg 9120
gaagtatctc ataccaancc ttaccgaaat aacctggatg gtatttatcc atgttaattc 9180
tgtggtgatg ttgaccaccg gccatacctc taccaccggg gtgctttctg tgcttaccga 9240
tacgaccttt accggctgag acgtgacctc tgtgctttct agtcttagtg aatctggaag 9300
gcattcttga ttagttggat gattgttctg ggatttaatg caaaaaaatc actaagaagg 9360
aaaaaaatca acggagaaag caaacgccat cttaaatata cgggatacag atgaaaggtt 9420
tgaacctatc tgggaaaata cgcattaaac aagcgaaaaa ctgcgaggaa aattgtttgc 9480
gtctctgcgg gctattcacg cgccagagga aaataggaaa aataacaggg cattagaaaa 9540
ataattttga ttttggtaat gtgtgggtcc ctggtgtaca gatgttacat tggttacagt 9600
actcttgttt ttgctgtgtt tttcgatgaa tctccaaaat ggttgttagc acatggaaga 9660
gtcaccgatg ctaagttatc tctatgtaag ctacgtggcg tgacttttga tgaagccgca 9720
caagagatac aggattggca actgcaaat                                   9749
```

<210> 93
<211> 1415
<212> PRT
<213> Caenorhabditis elegans

<400> 93

```
Met Ile Leu Arg His Phe Gly Phe Ile Leu Ile Ile Leu Thr Val Tyr
 1               5                  10                  15

Leu Ser Ser Thr His Ala Thr Thr Arg Lys His His Arg Arg Ser Ile
```

20 25 30

Asp Asn Asp Ala Arg Asn Leu Gly Phe Gln Phe Val Met Glu Pro Arg
35 40 45

Arg Asn Val Thr Val Met Glu Ser Ser Ser His Leu Leu Glu Cys Ser
50 55 60

Tyr Val Leu Ala His Glu Arg Leu Val His Asp Val Arg Ile Glu Trp
65 70 75 80

Lys Arg Asp Gly Val Leu Leu Ser Glu Arg Thr Ser Ser Arg Ile Lys
85 90 95

Val Met Ser Asn Gly Ser Leu Trp Ile Glu Ser Val Ser Ser Ala Glu
100 105 110

Glu Gly Thr Tyr Gln Cys Ala Val His Val Thr Thr Lys Ser Asp Gln
115 120 125

Thr Ser Asp Thr Trp Thr Phe Leu Ser Arg Lys Ala Thr Leu Arg Leu
130 135 140

Ala Asp Leu Ala Lys Phe Glu Leu Gln Ala Ile Asp Arg Thr Leu Ala
145 150 155 160

Lys Gly Gln Pro Thr Ala Phe His Cys Leu Ile Asn Ser Lys Pro Thr
165 170 175

Pro Thr Ala Val Trp Leu His Asn Asp Glu Pro Ile Val Asn Gly Gly
180 185 190

Glu Tyr His Ile Leu Pro Val Ser Asn Thr Leu Glu Ile Ser Ser Thr
195 200 205

Gln Ser Arg His Glu Gly Thr Tyr Arg Cys Thr Val Glu Gly Ala Gly
210 215 220

Lys Arg Arg Ser Ser Gln Thr Ala Arg Leu Thr Val Thr Thr Glu Thr
225 230 235 240

Val Ser Asn Glu Leu Val Phe Ile Thr Thr Pro Arg Leu Gln Val Val
245 250 255

Glu Gln Gly Asp Glu Phe Leu Leu Glu Cys Leu Val Ala Ser Leu Ile
260 265 270

Arg Pro Gln Val Arg Trp Leu Lys Asp Ser Arg Gln Ile Ile Val Asp
275 280 285

Gly Val Arg Ile Arg Arg Val Gly Val Ser Ser Ile Leu Val Ser Arg
290 295 300

Ala Ser Ile Glu Asp Thr Gly Leu Tyr Thr Cys Arg Ala Ser Asn Asn
305 310 315 320

Asp Asp Ser Ile Asp Arg Ala Val Ser Val Glu Val Arg Ala Pro Pro
325 330 335

```
Arg Ile Thr Thr Arg Pro Thr Thr Lys Val Ala Val Glu Thr Ala Asp
            340                 345                 350

Val Glu Leu Glu Cys Gly Thr Ala Ala Ala Arg Pro Glu Ala Arg Val
        355                 360                 365

Asn Trp Tyr Lys Asn Gly Glu Ala Ile Ile Gly Ser Glu Tyr Phe Val
    370                 375                 380

Ile Glu Pro Asn Arg Leu Arg Ile Leu Gly Val Val Arg Ala Asp Gln
385                 390                 395                 400

Ala Ile Tyr Gln Cys Ile Ala Glu Asn Asp Val Gly Ser Glu Gln Ala
                405                 410                 415

Ser Ala Gln Leu Leu Val Asp Ala Pro Asp Ser Ser Ser Val Ala Ala
            420                 425                 430

Ser Ser Gly Val Pro Met Thr Ser Ser Ala Pro Leu Gly Leu Arg Ser
        435                 440                 445

Thr Ser Ser Gly Ser Arg Phe Ile Asn Val Glu Trp Asp Pro Pro Val
    450                 455                 460

Gln Arg Asn Gly Asn Ile Met Arg Tyr His Ile Phe Tyr Lys Asp Asn
465                 470                 475                 480

Leu Ile Asp Arg Glu Arg Met Ile Asn Ser Ser Ser Thr Ser Ala Thr
                485                 490                 495

Leu Thr Ser Leu Gln Pro Ser Thr Met Tyr Leu Ile Arg Val Thr Ala
            500                 505                 510

Glu Asn Glu Ala Gly Met Gly Lys Phe Ser Asp Ser Leu Lys Val Thr
        515                 520                 525

Thr Asn Lys Glu Gln Ala Val Pro Gly Lys Val Ala Ser Leu Thr Thr
    530                 535                 540

Thr Ala Thr Gly Pro Glu Thr Ile Asp Ile Arg Trp Ser Pro Pro Ser
545                 550                 555                 560

Gly Gly Gln Pro Ala Leu Arg Tyr Lys Ile Phe Tyr Ser His Asp Pro
                565                 570                 575

Leu Glu Lys Asn Glu Lys Glu Thr Leu Ile Thr Thr Ser Thr Thr His
            580                 585                 590

Tyr Thr Leu His Gly Met Asp Lys Tyr Thr Gly Tyr Gln Ile Arg Ile
        595                 600                 605

Glu Ala Glu Gly Ser Asn Gly Ser Gly Leu Ser Ser Asp Thr Val Lys
    610                 615                 620

Val Arg Thr Gln Ser Asp Glu Pro Ser Ala Pro Pro Val Asn Ile Gln
625                 630                 635                 640
```

Ala Glu Ala Asp Ser Ser Thr Ser Val Arg Val Ser Trp Asp Glu Pro
645 650 655

Glu Glu Glu Ser Val Asn Gly Glu Ile Thr Gly Tyr Arg Leu Lys Tyr
660 665 670

Lys Thr Lys Ala Arg Gly Ala Lys Gly Asn Thr Leu Val Ile Asp Ala
675 680 685

Thr Ala Arg Glu Tyr Thr Met Gly Asn Leu Glu Pro Asn Thr Gln Tyr
690 695 700

Leu Ile Arg Met Ala Val Val Asn His Asn Gly Thr Gly Pro Phe Ser
705 710 715 720

Asp Trp Val Ser Ile Asp Thr Pro Gly Gln Asp Lys Glu Glu Arg Thr
725 730 735

Leu Gly Ala Pro Arg Glu Ile Arg Pro His Ala Gly Ile Asp Tyr Ile
740 745 750

Leu Val Ser Trp Leu Pro Pro Ala Asp Glu Gln Asn Leu Val Arg Gly
755 760 765

Tyr Gln Ile Gly Trp Gly Leu Ser Val Pro Asp Thr Glu Thr Ile Arg
770 775 780

Val Thr Ala Ser Thr Thr Gln Tyr Lys Ile Ala Arg Leu His Ser Glu
785 790 795 800

Arg Asp Tyr Val Ile Ser Leu Arg Ala Phe Asn Asn Leu Gly Ser Gly
805 810 815

Phe Pro Ile Tyr Glu Thr Val Arg Thr Leu Ser Arg Glu Thr Pro Ser
820 825 830

His Phe Asn Glu Asp Ser Asp Ser Asp Asp Ser Asp Val Gly Ser Ser
835 840 845

Glu Ser Thr Pro Val Gly Val Arg Ala Glu Ala Ile Ser Ala Thr Ser
850 855 860

Ile Arg Val Met Trp Thr Glu Ser Asp Glu Thr Ala Phe Asn Thr Gln
865 870 875 880

Tyr Thr Val Arg Tyr Ser Thr Ala Val Asp Gly Asn Gln His Arg Tyr
885 890 895

Val Asn Ser Thr Glu Thr Trp Ala Thr Val Glu Gly Leu Arg Pro Ala
900 905 910

Thr Glu Tyr Glu Phe Ala Val Arg Ala Val Ala Ser Asn Gly Gln Leu
915 920 925

Ser Thr Trp Ser Met Ala Thr Arg Asn Arg Thr Leu Ala Ala Pro Pro
930 935 940

Ser Ser Ala Pro Arg Asp Leu Thr Val Leu Pro Ala Glu Ser Gly Asp

945 950 955 960

Pro His Ser Ser Ser Leu His Trp Gln Pro Pro Lys Tyr Ser Asn Gly
965 970 975

Glu Ile Glu Glu Tyr Leu Val Phe Tyr Thr Asp Arg Ala Ser Leu Ala
980 985 990

Asp Lys Asp Trp Thr Ile Asn Tyr Val Ala Gly Asp Lys Leu Ser His
995 1000 1005

Gln Val Ser Asn Leu Leu Pro Lys Ala Asn Tyr Phe Phe Lys Ile Gln
1010 1015 1020

Ala Arg Asn Glu Lys Gly His Gly Pro Phe Ser Ser Val Val Gly Tyr
1025 1030 1035 1040

Thr Pro Ser Gly Gly Ala Ile Leu Ser Gly Lys Asp Arg His Asn Ala
1045 1050 1055

Arg Gly His Gly Ser Ala Ala Ser Gly Asp Thr Val Ser Leu Val Asp
1060 1065 1070

Gln Leu Gln Ser Leu Leu His Ser Asn Pro Leu Tyr Leu Ile Leu Leu
1075 1080 1085

Ala Ala Phe Ala Leu Ile Leu Ile Leu Thr Leu Ile Leu Ile Ile Met
1090 1095 1100

Cys Cys Trp Lys Arg Ser Ser Gly Gly Gly Arg Lys Asn Gly Tyr Gln
1105 1110 1115 1120

Ser Gly Lys Lys Thr Ser Ala Gly Ala Gly Ser Gly Gly Gly Ile Gly
1125 1130 1135

Gly Leu Gly Gly Pro Pro Asn Asp Leu Trp Ile Asn Gly Thr Gly Ser
1140 1145 1150

His Met Arg Ala Gly Ala Ser Asp Tyr Met Val Asp Gly Leu Ala Thr
1155 1160 1165

Ala His Leu Thr Ala Ala Asp Ile Glu Ser Pro Thr Pro Arg Tyr His
1170 1175 1180

His Leu Gln Gly Gln Gly Thr Leu Thr Arg Ser Tyr His Gln Ser Ser
1185 1190 1195 1200

Gln Ser Leu Glu Gly Arg Gln Arg Thr Pro Gln Val Val Tyr Thr Gly
1205 1210 1215

Thr Gly Arg His Gln Pro Ile Gln Arg Ile Asp Phe Glu Ser Pro Tyr
1220 1225 1230

Gly Ser Ser Ser Ala Ile Gly Ser Ala Ser Thr Pro Pro Leu Pro Met
1235 1240 1245

Gln Ala Pro Pro Ser Gly Pro Pro Thr Val Ile Asp Gly Tyr Arg Thr
1250 1255 1260

Leu Arg Gly Thr Pro Pro Asn Ser Ser Ala Ala Asn Ala Leu Arg Ser
1265 1270 1275 1280

Phe Thr Gln Leu Ala Gly Ala Thr Pro Pro Pro Pro His Ser Ala Ala
1285 1290 1295

Ser Ser Ser Ser Arg Pro Thr Ile Ile Ala Ala Gly Gly Arg Gln Val
1300 1305 1310

Pro Val Gly Arg Ala Thr Ala Gln Pro Arg Val Asn Val Ala Asn Ile
1315 1320 1325

Tyr Ser Pro Phe Ala Ser Cys Ser Ala Ser Ser Asp Ala Gly Glu Ser
1330 1335 1340

Asp Lys Lys Ser Gly Glu Cys Met Glu Met Arg Glu Thr Thr Pro Ile
1345 1350 1355 1360

Lys Ser Asn Thr Ala Gly Ser Ser Asn Gly Glu Lys Met Asn Thr Asn
1365 1370 1375

Met Asn Pro Ser His Ser Ala Glu Asp Leu Asn Ala His Leu Glu Asn
1380 1385 1390

Leu Asp Thr Met Leu Asp Asp Leu Gln Gln Leu Gln His Asn Leu His
1395 1400 1405

Phe Glu Thr Ser Met Asp Lys
1410 1415

<210> 94
<211> 4248
<212> DNA
<213> Caenorhabditis elegans

<400> 94
atgattttgc gacatttcgg tttcattctt ataattctga cagtttatct gtcatctact 60
catgccacaa cacggaaaca tcatcgtcga tcgattgata atgatgctcg gaatcttggc 120
tttcaattcg taatggaacc acgacggaat gtgactgtta tggaatcatc gtctcattta 180
ttagaatgta gttatgttct ggctcatgag agacttgtac acgatgtccg aattgaatgg 240
aaacgtgatg gtgtgttgct cagtgaacgg acatcttcta gaataaaagt aatgtcaaat 300
ggttcattat ggatagaatc agtgagcagt gccgaagaag gaacatatca atgcgctgta 360
catgtgacta caaaaagcga tcaaacaagt gatacgtgga cattttttgag tcgaaaagct 420
acgttacgat tggcggattt ggcaaagttt gaactgcaag cgattgatcg aactctagca 480
aaagggcagc caactgcgtt tcattgtctt attaactcga aaccaacccc aacagcagtt 540
tggcttcaca acgatgaacc tattgtaaat ggtggcgaat atcacattct accagtgtca 600
aacactttgg agattagttc cactcaatct agacacgaag ggacgtatag gtgtacggta 660
gaaggtgctg ggaagagaag aagtagccaa actgcacgat tgaccgttac tactgaaaca 720
gtatcgaatg agctagtttt catcacaacg ccacgtctac aagttgttga acaaggcgac 780
gaattcttac ttgaatgcct tgttgcctca ctaatcagac ctcaagttag atggctcaaa 840
gattctcgtc aaattattgt cgatggagtt cgtattcgac gtgttggcgt atcgtcaatt 900
cttgtctcac gtgcttctat cgaagatact gggctataca cttgtcgagc ttcgaataat 960
gatgattcaa ttgatcgagc tgtatctgta gaagttcgtg ctccgccgag aatcactaca 1020
cgacctacca cgaaagttgc tgttgaaact gctgatgttg aattagagtg tggaacggcg 1080
gctgctagac cagaggcacg cgttaattgg tacaaaaatg gagaagcaat aattgggagc 1140
gagtattttg taattgagcc gaacagactc cgcatccttg gagtcgtacg agccgatcaa 1200

gcaatctacc aatgtattgc agaaaatgat gttggaagtg aacaagctag tgcacaactt 1260
cttgttgatg ctccagattc atcatcagtg gctgcttcct caggagttcc aatgacttcg 1320
tctgcaccgt taggcctccg ttcgacgtct tctggttctc gctttataaa tgttgaatgg 1380
gatccaccag ttcaaagaaa tgggaatatt atgagatatc atatattcta caaagataat 1440
ttaatcgaca gagaaaggat gatcaattca tcaagtacat ctgcaacgtt aacatctcta 1500
caaccatcaa caatgtattt aattcgagtt actgctgaaa atgaagctgg aatgggaaag 1560
tttagtgatt cacttaaagt taccacaaac aaagaacaag cagttcctgg aaaagtcgca 1620
tcactaacaa caactgcaac tggccccgaa acgattgaca tccgttggag cccaccaagt 1680
ggaggacaac ctgcccttcg atacaaaata ttctattcgc atgatccatt ggagaaaaat 1740
gaaaaagaaa cattgattac gacctcgaca acacattata cacttcatgg aatggataaa 1800
tacactggat atcagattcg aattgaagct gaaggatcca atggatcggg gctttcaagt 1860
catactgtga aagtacgaac acaatccgat gaaccatcag caccgcctgt aaacattcaa 1920
gcagaagctg atagctcaac aagtgttcga gttagttggg atgagccaga agaggaatct 1980
gttaacggag aaattactgg atatcggttg aagtacaaga caaaagctcg tggagccaaa 2040
ggcaatactc ttgtgattga cgccactgct agagagtata ctatgggaaa tttggaaccg 2100
aatacacagt atctgataag aatggcagta gtgaatcaca atggaacggg acccttctcg 2160
gattgggttt ctattgatac tcctggacaa gataaagaag aacgaacact tggagcaccg 2220
agagagattc gaccacatgc tggaattgat tatattcttg tatcatggct tccaccagca 2280
gatgaacaga atcttgttag aggatatcaa attggatggg gattgagtgt tcctgatact 2340
gaaacaattc gtgtaactgc ttccacaact caatataaaa ttgccagatt gcattcggaa 2400
cgagattatg tgatttctct tcgagcattt aataatctag gatctggatt tccgatttat 2460
gagactgtcc gaacattgtc tcgtgaaact ccatcacatt tcaatgaaga ttctgattct 2520
gacgattctg acgttggatc ctcagaatct actccagttg gagtgcgtgc tgaggcaatc 2580
tctgcaacat caatacgtgt tatgtggacg gaatctgatg aaactgcttt caatacacag 2640
tatacagttc gatatagtac agctgttgat gggaatcaac atagatacgt gaatagcacc 2700
gagacatggg ctactgtgga aggactacga ccagctactg agtatgagtt tgcggttaga 2760
gctgttgcat caaatgggca gttgagtact tggtcaatgg ctaccaggaa tcgaacatta 2820
gcggctccac caagttcagc tccacgtgat ctcactgtgc tgcctgcaga aagtggcgat 2880
cctcactctt caagtcttca ttggcaacct ccgaaatatt caaatggaga aattgaagaa 2940
tatttagtgt tttatacgga tcgagcttct cttgcggaca aagattggac aattaactat 3000
gtagctggtg ataagttgtc acatcaagtc agcaatttac ttccaaaagc taattatttc 3060
ttcaagattc aagcaagaaa tgagaaaggt catggcccat tctcttctgt agttggatac 3120
actccatcag gcggtgctat tctatcagga aaagatcgtc acaatgcacg tgggcatgga 3180
tcagcggctt ctggcgatac agtttcctta gtcgatcaac ttcaatctct tcttcactct 3240
aacccattgt atctcatact tttggctgcg tttgctctga ttctgatcct tacactgatt 3300
cttatcataa tgtgttgttg gaaacgatct agtggtggag ggcgaaagaa tggatatcag 3360
tctggaaaaa agacgagtgc cggagctgga agtggaggtg gaataggtgg tcttggaggt 3420
ccacctaatg atttatggat taatggaact ggaagtcata tgagagctgg tgcttctgat 3480
tatatggttg atggtctagc aaccgcccat cttactgctg ctgacattga atcacccaca 3540
cctagatacc atcatctaca agggcaagga actctaacac gaagttatca tcaaagctca 3600
caaagtcttg aaggacgaca acgaacacct caagtggttt atacaggaac tggaaggcat 3660
cagccgattc aaagaatcga ctttgaaagt ccttatggtt catcatctgc aattggatcc 3720
gcctcgactc cgcctctccc aatgcaagct ccaccttctg gtccacctac agtaatcgat 3780
ggataccgta ctctacgtgg aacacctcca aattcatcag ctgcaaatgc tcttcggtca 3840
ttcactcaat tggcaggagc aacaccacct ccaccacact ctgcggcatc atcatcttct 3900
cgaccgacta ttattgcagc cggtggacgg caagttccag ttggaagagc aactgctcaa 3960
ccacgtgtca atgttgcaaa tatctacagt ccatttgcat catgctctgc atcttctgat 4020
gcaggagaat cagataaaaa gagtggagaa tgtatggaaa tgagagaaac cacacccatt 4080
aaatctaata cagctggttc atcgaatgga gagaagatga acacaaatat gaatccttca 4140
cactcggcag aagatctcaa cgcacatctt gagaaccttg acacaatgct tgatgatcta 4200
caacaattac agcacaattt gcattttgag acgagtatgg ataagtaa 4248

<210> 95
<211> 1447
<212> PRT
<213> Homo sapiens

<400> 95

```
Met Glu Asn Ser Leu Arg Cys Val Trp Val Pro Lys Leu Ala Phe Val
 1               5                  10                  15

Leu Phe Gly Ala Ser Leu Leu Ser Ala His Leu Gln Val Thr Gly Phe
            20                  25                  30

Gln Ile Lys Ala Phe Thr Ala Leu Arg Phe Leu Ser Glu Pro Ser Asp
        35                  40                  45

Ala Val Thr Met Arg Gly Gly Asn Val Leu Leu Asp Cys Ser Ala Glu
    50                  55                  60

Ser Asp Arg Gly Val Pro Val Ile Lys Trp Lys Lys Asp Gly Ile His
65                  70                  75                  80

Leu Ala Leu Gly Met Asp Glu Arg Lys Gln Gln Leu Ser Asn Gly Ser
                85                  90                  95

Leu Leu Ile Gln Asn Ile Leu His Ser Arg His His Lys Pro Asp Glu
            100                 105                 110

Gly Leu Tyr Gln Cys Glu Ala Ser Leu Gly Asp Ser Gly Ser Ile Ile
        115                 120                 125

Ser Arg Thr Ala Lys Val Ala Val Ala Gly Pro Leu Arg Phe Leu Ser
    130                 135                 140

Gln Thr Glu Ser Val Thr Ala Phe Met Gly Asp Thr Val Leu Leu Lys
145                 150                 155                 160

Cys Glu Val Ile Gly Glu Pro Met Pro Thr Ile His Trp Gln Lys Asn
                165                 170                 175

Gln Gln Asp Leu Thr Pro Ile Pro Gly Asp Ser Arg Val Val Val Leu
            180                 185                 190

Pro Ser Gly Ala Leu Gln Ile Ser Arg Leu Gln Pro Gly Asp Ile Gly
        195                 200                 205

Ile Tyr Arg Cys Ser Ala Arg Asn Pro Ala Ser Ser Arg Thr Gly Asn
    210                 215                 220

Glu Ala Glu Val Arg Ile Leu Ser Asp Pro Gly Leu His Arg Gln Leu
225                 230                 235                 240

Tyr Phe Leu Gln Arg Pro Ser Asn Val Val Ala Ile Glu Gly Lys Asp
                245                 250                 255

Ala Val Leu Glu Cys Cys Val Ser Gly Tyr Pro Pro Pro Ser Phe Thr
            260                 265                 270

Trp Leu Arg Gly Glu Glu Val Ile Gln Leu Arg Ser Lys Lys Tyr Ser
        275                 280                 285

Leu Leu Gly Gly Ser Asn Leu Leu Ile Ser Asn Val Thr Asp Asp Asp
    290                 295                 300

Ser Gly Met Tyr Thr Cys Val Val Thr Tyr Lys Asn Glu Asn Ile Ser
```

305 310 315 320

Ala Ser Ala Glu Leu Thr Val Leu Val Pro Pro Trp Phe Leu Asn His
325 330 335

Pro Ser Asn Leu Tyr Ala Tyr Glu Ser Met Asp Ile Glu Phe Glu Cys
340 345 350

Thr Val Ser Gly Lys Pro Val Pro Thr Val Asn Trp Met Lys Asn Gly
355 360 365

Asp Val Val Ile Pro Ser Asp Tyr Phe Gln Ile Val Gly Gly Ser Asn
370 375 380

Leu Arg Ile Leu Gly Val Val Lys Ser Asp Glu Gly Phe Tyr Gln Cys
385 390 395 400

Val Ala Glu Asn Glu Ala Gly Asn Ala Gln Thr Ser Ala Gln Leu Ile
405 410 415

Val Pro Lys Pro Ala Ile Pro Ser Ser Ser Val Leu Pro Ser Ala Pro
420 425 430

Arg Asp Val Val Pro Val Leu Val Ser Ser Arg Phe Val Arg Leu Ser
435 440 445

Trp Arg Pro Pro Ala Glu Ala Lys Gly Asn Ile Gln Thr Phe Thr Val
450 455 460

Phe Phe Ser Arg Glu Gly Asp Asn Arg Glu Arg Ala Leu Asn Thr Thr
465 470 475 480

Gln Pro Gly Ser Leu Gln Leu Thr Val Gly Asn Leu Lys Pro Glu Ala
485 490 495

Met Tyr Thr Phe Arg Val Val Ala Tyr Asn Glu Trp Gly Pro Gly Glu
500 505 510

Ser Ser Gln Pro Ile Lys Val Ala Thr Gln Pro Glu Leu Gln Val Pro
515 520 525

Gly Pro Val Glu Asn Leu Gln Ala Val Ser Thr Ser Pro Thr Ser Ile
530 535 540

Leu Ile Thr Trp Glu Pro Pro Ala Tyr Ala Asn Gly Pro Val Gln Gly
545 550 555 560

Tyr Arg Leu Phe Cys Thr Glu Val Ser Thr Gly Lys Glu Gln Asn Ile
565 570 575

Glu Val Asp Gly Leu Ser Tyr Lys Leu Glu Gly Leu Lys Lys Phe Thr
580 585 590

Glu Tyr Ser Leu Arg Phe Leu Ala Tyr Asn Arg Tyr Gly Pro Gly Val
595 600 605

Ser Thr Asp Asp Ile Thr Val Val Thr Leu Ser Asp Val Pro Ser Ala
610 615 620

Pro Pro Gln Asn Val Ser Leu Glu Val Val Asn Ser Arg Ser Ile Lys
625 630 635 640

Val Ser Trp Leu Pro Pro Pro Ser Gly Thr Gln Asn Gly Phe Ile Thr
645 650 655

Gly Tyr Lys Ile Arg His Arg Lys Thr Thr Arg Arg Gly Glu Met Glu
660 665 670

Thr Leu Glu Pro Asn Asn Leu Trp Tyr Leu Phe Thr Gly Leu Glu Lys
675 680 685

Gly Ser Gln Tyr Ser Phe Gln Val Ser Ala Met Thr Val Asn Gly Thr
690 695 700

Gly Pro Pro Ser Asn Trp Tyr Thr Ala Glu Thr Pro Glu Asn Asp Leu
705 710 715 720

Asp Glu Ser Gln Val Pro Asp Gln Pro Ser Ser Leu His Val Arg Pro
725 730 735

Gln Thr Asn Cys Ile Ile Met Ser Trp Thr Pro Pro Leu Asn Pro Asn
740 745 750

Ile Val Val Arg Gly Tyr Ile Ile Gly Tyr Gly Val Gly Ser Pro Tyr
755 760 765

Ala Glu Thr Val Arg Val Asp Ser Lys Gln Arg Tyr Tyr Ser Ile Glu
770 775 780

Arg Leu Glu Ser Ser Ser His Tyr Val Ile Ser Leu Lys Ala Phe Asn
785 790 795 800

Asn Ala Gly Glu Gly Val Pro Leu Tyr Glu Ser Ala Thr Thr Arg Ser
805 810 815

Ile Thr Asp Pro Thr Asp Pro Val Asp Tyr Tyr Pro Leu Leu Asp Asp
820 825 830

Phe Pro Thr Ser Val Pro Asp Leu Ser Thr Pro Met Leu Pro Pro Val
835 840 845

Gly Val Gln Ala Val Ala Leu Thr His Asp Ala Val Arg Val Ser Trp
850 855 860

Ala Asp Asn Ser Val Pro Lys Asn Gln Lys Thr Ser Glu Val Arg Leu
865 870 875 880

Tyr Thr Val Arg Trp Arg Thr Ser Phe Ser Ala Ser Ala Lys Tyr Lys
885 890 895

Ser Glu Asp Thr Thr Ser Leu Ser Tyr Thr Ala Thr Gly Leu Lys Pro
900 905 910

Asn Thr Met Tyr Glu Phe Ser Val Met Val Thr Lys Asn Arg Arg Ser
915 920 925

Ser Thr Trp Ser Met Thr Ala His Ala Thr Thr Tyr Glu Ala Ala Pro
930 935 940

Thr Ser Ala Pro Lys Asp Phe Thr Val Ile Thr Arg Glu Gly Lys Pro
945 950 955 960

Arg Ala Val Ile Val Ser Trp Gln Pro Pro Leu Glu Ala Asn Gly Lys
965 970 975

Ile Thr Ala Tyr Ile Leu Phe Tyr Thr Leu Asp Lys Asn Ile Pro Ile
980 985 990

Asp Asp Trp Ile Met Glu Thr Ile Ser Gly Asp Arg Leu Thr His Gln
995 1000 1005

Ile Met Asp Leu Asn Leu Asp Thr Met Tyr Tyr Phe Arg Ile Gln Ala
1010 1015 1020

Arg Asn Ser Lys Gly Val Gly Pro Leu Ser Asp Pro Ile Leu Phe Arg
1025 1030 1035 1040

Thr Leu Lys Val Glu His Pro Asp Lys Met Ala Asn Asp Gln Gly Arg
1045 1050 1055

His Gly Asp Gly Gly Tyr Trp Pro Val Asp Thr Asn Leu Ile Asp Arg
1060 1065 1070

Ser Thr Leu Asn Glu Pro Pro Ile Gly Gln Met His Pro Pro His Gly
1075 1080 1085

Ser Val Thr Pro Gln Lys Asn Ser Asn Leu Leu Val Ile Ile Val Val
1090 1095 1100

Thr Val Gly Val Ile Thr Val Leu Val Val Val Ile Val Ala Val Ile
1105 1110 1115 1120

Cys Thr Arg Arg Ser Ser Ala Gln Gln Arg Lys Lys Arg Ala Thr His
1125 1130 1135

Ser Ala Gly Lys Arg Lys Gly Ser Gln Lys Asp Leu Arg Pro Pro Asp
1140 1145 1150

Leu Trp Ile His His Glu Glu Met Glu Met Lys Asn Ile Glu Lys Pro
1155 1160 1165

Ser Gly Thr Asp Pro Ala Gly Arg Asp Ser Pro Ile Gln Ser Cys Gln
1170 1175 1180

Asp Leu Thr Pro Val Ser His Ser Gln Ser Glu Thr Gln Leu Gly Ser
1185 1190 1195 1200

Lys Ser Thr Ser His Ser Gly Gln Asp Thr Glu Glu Ala Gly Ser Ser
1205 1210 1215

Met Ser Thr Leu Glu Arg Ser Leu Ala Ala Arg Arg Ala Pro Arg Ala
1220 1225 1230

Lys Leu Met Ile Pro Met Asp Ala Gln Ser Asn Asn Pro Ala Val Val

1235 1240 1245

Ser Ala Ile Pro Val Pro Thr Leu Glu Ser Ala Gln Tyr Pro Gly Ile
1250 1255 1260

Leu Pro Ser Pro Thr Cys Gly Tyr Pro His Pro Gln Phe Thr Leu Arg
1265 1270 1275 1280

Pro Val Pro Phe Pro Thr Leu Ser Val Asp Arg Gly Phe Gly Ala Gly
1285 1290 1295

Arg Ser Gln Ser Val Ser Glu Gly Pro Thr Thr Gln Gln Pro Pro Met
1300 1305 1310

Leu Pro Pro Ser Gln Pro Glu His Ser Ser Ser Glu Glu Ala Pro Ser
1315 1320 1325

Arg Thr Ile Pro Thr Ala Cys Val Arg Pro Thr His Pro Leu Arg Ser
1330 1335 1340

Phe Ala Asn Pro Leu Leu Pro Pro Pro Met Ser Ala Ile Glu Pro Lys
1345 1350 1355 1360

Val Pro Tyr Thr Pro Leu Leu Ser Gln Pro Gly Pro Thr Leu Pro Lys
1365 1370 1375

Thr His Val Lys Thr Ala Ser Leu Gly Leu Ala Gly Lys Ala Arg Ser
1380 1385 1390

Pro Leu Leu Pro Val Ser Val Pro Thr Ala Pro Glu Val Ser Glu Glu
1395 1400 1405

Ser His Lys Pro Thr Glu Asp Ser Ala Asn Val Tyr Glu Gln Asp Asp
1410 1415 1420

Leu Ser Glu Gln Met Ala Ser Leu Glu Gly Leu Met Lys Gln Leu Asn
1425 1430 1435 1440

Ala Ile Thr Gly Ser Ala Phe
1445

<210> 96
<211> 4344
<212> DNA
<213> Homo sapiens

<400> 96
atggagaata gtcttagatg tgtttgggta cccaagctgg cttttgtact cttcggagct 60
tccttgctca gcgcgcatct tcaagtaacc ggttttcaaa ttaaagcttt cacagcactg 120
cgcttcctct cagaaccttc tgatgccgtc acaatgcggg gaggaaatgt cctcctcgac 180
tgctccgcgg agtccgaccg aggagttcca gtgatcaagt ggaagaaaga tggcattcat 240
ctggccttgg gaatggatga aaggaagcag caactttcaa atgggtctct gctgatacaa 300
aacatacttc attccagaca ccacaagcca gatgagggac tttaccaatg tgaggcatct 360
ttaggagatt ctggctcaat tattagtcgg acagcaaaag ttgcagtagc aggaccactg 420
aggttccttt cacagacaga atctgtcaca gccttcatgg gagacacagt gctactcaag 480
tgtgaagtca ttggggagcc catgccaaca atccactggc agaagaacca acaagacctg 540
actccaatcc caggtgactc ccgagtggtg gtcttgccct ctggagcatt gcagatcagc 600

cgactccaac cgggggacat tggaatttac cgatgctcag ctcgaaatcc agccagctca 660
agaacaggaa atgaagcaga agtcagaatt ttatcagatc caggactgca tagacagctg 720
tattttctgc aaagaccatc caatgtagta gccattgaag gaaaagatgc tgtcctggaa 780
tgttgtgttt ctggctatcc tccaccaagt tttacctggt tacgaggcga ggaagtcatc 840
caactcaggt ctaaaaagta ttctttattg ggtggaagca acttgcttat ctccaatgtg 900
acagatgatg acagtggaat gtatacctgt gttgtcacat ataaaaatga gaatattagt 960
gcctctgcag agctcacagt cttggttccg ccatggtttt taaatcatcc ttccaacctg 1020
tatgcctatg aaagcatgga tattgagttt gaatgtacag tctctggaaa gcctgtgccc 1080
actgtgaatt ggatgaagaa tggagatgtg gtcattccta gtgattattt tcagatagtg 1140
ggaggaagca acttacggat acttggggtg gtgaagtcag atgaaggctt ttatcaatgt 1200
gtggctgaaa atgaggctgg aaatgcccag accagtgcac agctcattgt ccctaagcct 1260
gcaatcccaa gctccagtgt cctcccttcg gctcccagag atgtggtccc tgtcttggtt 1320
tccagccgat ttgtccgtct cagctggcgc ccacctgcag aagcgaaagg gaacattcaa 1380
actttcacgg tctttttctc cagagaaggt gacaacaggg aacgagcatt gaatacaaca 1440
cagcctgggt cccttcagct cactgtggga aacctgaagc cagaagccat gtacaccttt 1500
cgagttgtgg cttacaatga atggggaccg ggagagagtt ctcaacccat caaggtggcc 1560
acacagcctg agttgcaagt tccagggcca gtagaaaacc tgcaagctgt atctacctca 1620
cctacctcaa ttcttattac ctgggaaccc cctgcctatg caaacggtcc agtccaaggt 1680
tacagattgt tctgcactga ggtgtccaca ggaaaagaac agaatataga ggttgatgga 1740
ctatcttata aactggaagg cctgaaaaaa ttcaccgaat atagtcttcg attcttagct 1800
tataatcgct atggtccggg cgtctctact gatgatataa cagtggttac actttctgac 1860
gtgccaagtg ccccgcctca gaacgtctcc ctggaagtgg tcaattcaag aagtatcaaa 1920
gttagctggc tgcctcctcc atcaggaaca caaaatggat ttattaccgg ctataaaatt 1980
cgacacagaa agacgacccg caggggtgag atggaaacac tggagccaaa caacctctgg 2040
tacctattca caggactgga gaaaggaagt cagtacagtt tccaggtgtc agccatgaca 2100
gtcaatggta ctggaccacc ttccaactgg tatactgcag agactccaga gaatgatcta 2160
gatgaatctc aagttcctga tcaaccaagc tctcttcatg tgaggcccca gactaactgc 2220
atcatcatga gttggactcc tcccttgaac ccaaacatcg tggtgcgagg ttatattatc 2280
ggttatggcg ttgggagccc ttacgctgag acagtgcgtg tggacagcaa gcagcgatat 2340
tattccattg agaggttaga gtcaagttcc cattatgtaa tctccctaaa agcttttaac 2400
aatgccggag aaggagttcc tctttatgaa agtgccacca ccaggtctat aaccgatccc 2460
actgacccag ttgattatta tcctttgctt gatgatttcc ccacctcggt cccagatctc 2520
tccacccccca tgctcccacc agtaggtgta caggctgtgg ctcttaccca tgatgctgtg 2580
agggtcagct gggcagacaa ctctgtccct aagaaccaaa agacgtctga ggtgcgactt 2640
tacaccgtcc ggtggagaac cagcttttct gcaagtgcaa aatacaagtc agaagacaca 2700
acatctctaa gttacacagc aacaggcctc aaaccaaaca caatgtatga attctcggtc 2760
atggtaacaa aaaacagaag gtccagtact tggagcatga ctgcacatgc caccacgtat 2820
gaagcagccc ccacctctgc tcccaaggac tttacagtca ttactaggga agggaagcct 2880
cgtgccgtca ttgtgagttg gcagcctccc ttggaagcca atgggaaaat tactgcttac 2940
atcttatttt ataccttgga caagaacatc ccaattgatg actggattat ggaaacaatc 3000
agtggtgata ggcttactca tcaaatcatg gatctcaacc ttgatactat gtattacttt 3060
cgaattcaag cacgaaattc aaaaggagtg ggccactct ctgatcccat cctcttcagg 3120
actctgaaag tggaacaccc tgacaaaatg gctaatgacc aaggtcgtca tggagatgga 3180
ggttattggc cagttgatac taatttgatt gatagaagca ccctaaatga gccgccaatt 3240
ggacaaatgc accccccgca tggcagtgtc actcctcaga agaacagcaa cctgcttgtg 3300
atcattgtgg tcaccgttgg tgtcatcaca gtgctggtag tggtcatcgt ggctgtgatt 3360
tgcacccgac gctcttcagc ccagcagaga aagaaacggg ccacccacag tgctggcaaa 3420
aggaagggca gccagaagga cctccgaccc cctgatcttt ggatccatca tgaagaaatg 3480
gagatgaaaa atattgaaaa gccatctggc actgaccctg caggaaggga ctctcccatc 3540
caaagttgcc aagacctcac accagtcagc cacagccagt cagaaaccca actgggaagc 3600
aaaagcacct ctcattcagg tcaagacact gaggaagcag ggagctctat gtccactctg 3660
gagaggtcgc tggctgcacg ccgagccccc cgggccaagc tcatgattcc catggatgcc 3720
cagtccaaca atcctgctgt cgtgagcgcc atcccggtgc caacgctaga aagtgcccag 3780
tacccaggaa tcctcccgtc tcccacctgt ggatatcccc acccgcagtt cactctccgg 3840
cctgtgccat tcccaacact ctcagtggac cgaggtttcg gagcaggaag aagtcagtca 3900
gtgagtgaag gaccaactac ccaacaacca cctatgctgc ccccatctca gcctgagcat 3960
tctagcagcg aggaggcacc aagcagaacc atccccacag cttgtgttcg accaactcac 4020
ccactccgca gctttgctaa tcctttgcta cctccaccaa tgagtgcaat agaaccgaaa 4080

# INTERNATIONAL SEARCH REPORT

| Internat ial Application No |
|---|
| PCT/EP 00/05108 |

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC 7 C12N15/12 C07K14/71 C12Q1/68 G01N33/50 G01N33/68
C07K16/18 C07K14/435

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EMBASE

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | ACKERMAN SUSAN L ET AL: "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23." GENOMICS, vol. 52, no. 2, 1998, pages 205-208, XP000946854 ISSN: 0888-7543 cited in the application | 3,9,15 |
| A | the whole document | 1-18 |
| | --- | |
| A | WO 98 37085 A (UNIV CALIFORNIA) 27 August 1998 (1998-08-27) the whole document | 1-28, 30-59, 61-64, 66,67,69 |
| | --- | |
| | -/-- | |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 17 October 2000 | 08. 01. 2001 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | ANDRES S.M. |

# INTERNATIONAL SEARCH REPORT

Intern.  ial Application No
PCT/EP 00/05108

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | WO 97 14424 A (UNIV CALIFORNIA) 24 April 1997 (1997-04-24) the whole document --- | 19-23 |
| A | COLAVITA ANTONIO ET AL: "Suppressors of ectopic UNC-5 growth cone steering identify eight genes involved in axon guidance in Caenorhabditis elegans." DEVELOPMENTAL BIOLOGY, vol. 194, no. 1, 1 February 1998 (1998-02-01), pages 72-85, XP000946782 ISSN: 0012-1606 cited in the application the whole document ----- | 23-25, 27,28 |

International application No.
PCT/EP 00/05108

# INTERNATIONAL SEARCH REPORT

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see further information sheet invention 1

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-18 (totally) and 19-28,30-59,61-64,66-67, 69 (all partially)

1.1. Claims: 1-6 (totally) and 19-28,30-59,61-64,66-67, 69 (all partially)

A human Unc-5Cb protein (SEQ ID 2), nucleic acids encoding it (SEQ ID 1), vectors and cells expressing it and an antibody binding thereto. Methods for identifying compounds which are capable of modulating the binding of Unc-5Cb to an interacting protein.

1.2. Claims: 7-12,71,85 (totally) and 19-28,30-59,61-64, 66-67,69 (all partially)

As for subject 1.1, but concerning a human Unc-5Cc protein (SEQ ID 4).

1.3. Claims: 13-18 (totally) and 19-28,30-59,61-64,66-67, 69 (all partially)

As for subject 1.1, but concerning a human Unc-5C8 protein (SEQ ID 6).

2. Claims: 19,29-58 (all partially)

A method, as characterised in claim 19, for identifying compounds capable of modulating the binding of an unc-5 protein to an interacting protein.

3. Claims: 20,29-58 (all partially)

A method, as characterised in claim 20, for identifying compounds capable of modulating the binding of an unc-5 protein to an interacting protein.

4. Claims: 21,29-58 (all partially)

A method, as characterised in claim 21, for identifying compounds capable of modulating the binding of an unc-5 protein to an interacting protein.

5. Claims: 22,29-58 (all partially)

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

A method, as characterised in claim 22, for identifying compounds capable of modulating the binding of an unc-5 protein to an interacting protein.

6. Claims: 23-58 (all partially)

A method, as characterised in claim 23, for identifying compounds capable of modulating the binding of an unc-5 protein to an interacting protein.

7. Claims: 59,61-64,66-67,69 (all partially) and claims 60,65, 68 (totally)

A method for identifying compounds reducing or inhibiting the lethal phenotype associated with the expression of an UNC-5 death domain in yeast.

8. Claims: 70,80-84 (totally) and 19-28,30-59,61-64,66-67, 69 (all partially)

A nucleic acid encoding the human unc-5H1 homolog (SEQ ID 7), probes and antisense nucleic acids hybridizing therewith, vectors and cells comprising it. Methods for identifying compounds which are capable of modulating the binding of Unc-5H1 to an interacting protein.

9. Claims: 72-73 (totally) and 19-31,53 (all partially)

A nucleic acid obtainable by digestion of pYMP17 with EcoRI and XhoI (SEQ ID 56). Methods for identifying compounds which are capable of modulating the binding of Unc-5 protein to the protein encoded by SEQ ID 56.

10. Claims: 74-75 (totally) and 19-31,54 (all partially)

A nucleic acid obtainable by digestion of pYMP6 with EcoRI and XhoI (SEQ ID 54). Methods for identifying compounds which are capable of modulating the binding of Unc-5 protein to the protein encoded by SEQ ID 54.

11. Claims: 76-77 (totally) and 19-31,57 (all partially)

A nucleic acid obtainable by digestion of pYMP11 with EcoRI and XhoI (SEQ ID 61). Methods for identifying compounds which are capable of modulating the binding of Unc-5 protein to the protein encoded by SEQ ID 61.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

12. Claims: 78-79 (totally) and 19-31,58 (all partially)

    A nucleic acid obtainable by digestion of pYMP12 with EcoRI and XhoI (SEQ ID 63). Methods for identifying compounds which are capable of modulating the binding of Unc-5 protein to the protein encoded by SEQ ID 63.

Please note that all inventions mentioned under item 1, although not necessarily linked by a common inventive concept, could be searched without effort justifying an additional fee.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 00/05108

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9837085 A | 27-08-1998 | US 5939271 A | 17-08-1999 |
| | | AU 718795 B | 20-04-2000 |
| | | AU 6174498 A | 09-09-1998 |
| | | EP 0973794 A | 26-01-2000 |
| WO 9714424 A | 24-04-1997 | US 5747262 A | 05-05-1998 |
| | | AU 688698 B | 12-03-1998 |
| | | AU 7433496 A | 07-05-1997 |
| | | CA 2207505 A | 24-04-1997 |
| | | EP 0802795 A | 29-10-1997 |
| | | JP 10511558 T | 10-11-1998 |
| | | US 6087326 A | 11-07-2000 |